Applications and Industry®

November 1961

Transactions Papers

General Applications Division

60-1028 Effect of Variable Plasma Conductivity—MHD Converter.....Coe, Eisen . . . 225

61-208 Economic Justification of Railway Electrification in U. S.....Cross . . . 232

Industry Division

61-802 Optimum Synthesis of Random Sampling Multipole Filters.....Hsieh . . . 239

61-743 Sampled-Data Control Systems with Transport Lag.....Šiljak . . . 247

61-819 Mathematical Models for Time-Domain Design.....Wang . . . 252

61-712 Signal Stabilization of Control System.....Oldenburger, Sridhar . . . 260

61-710 Frequency Response of Nonlinear Closed-Loop Systems.....McAllister . . . 268

61-709 Stabilization of Feedback Systems.....Mahalanabis . . . 277

60-1066 Thermoelectricity Application Considerations.....Sorensen . . . 285

61-713 Advances in Analysis and Synthesis of Nonlinear Systems.....Wolf . . . 289

61-752 Feedback Compensation: A Design Technique..Thaler, Bronzino, Kirk . . . 300

61-711 Can Electric Actuators Meet Missile Requirements?....Newton, Rasche . . . 306

Conference Papers Open for Discussion.....See 3rd Cover

NUMBER 57

Published Bimonthly by

AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS

Communication and Electronics—November 1961

Instrumentation Division

61-721 General Description of D-C Digital Voltmeters.....Stansbury . . . 465

Communication Division

61-827 Features of an Electronic Crosspoint PABX.....Van Bosse, DeCicco . . . 471
61-722 Centrex Service: New Design for Group Telephone Service.....Shea . . . 474
61-113 Pole-Zero Techniques Applied to V-F Telephone Lines.....Fleming . . . 482
61-181 Noise and Intermodulation in Closed-Circuit TV.....Collins, Williams . . . 486
61-825 Transmission Network of Electronic Crosspoint PABX.....Kowalik . . . 491
61-826 Logical Control of Electronic Crosspoint PABX.....Sanders . . . 496
61-822 Reliable Data Transmission Through Noisy Media.....Melas . . . 501

Science and Electronics Division

61-732 Large-Scale On-Line Data Processing Systems.....Levine . . . 505
60-1007 Computer Program for Preparing Wiring Diagrams..Kirby, Rosenthal . . . 509
61-150 Magnetically Regulated D-C to D-C Converter Power Supply.....Sager . . . 513
61-717 Regulated Power Supply Using Variable Output Oscillator.....Jackson . . . 518
61-814 Forcing Circuitry: Sequential Building Blocks for Logic.....Meade . . . 522
61-718 Inverter with Improved Commutation.....McMurray, Shattuck . . . 531
61-725 A Comparison of Computers.....Curl . . . 542
61-724 Results of Simulation Comparison of Control Computers.....Sendzuk . . . 547

Education Committee

61-733 Educating Electrical Engineers for Professional Careers.....Linke . . . 551
61-735 Motivation Through Challenge.....Johnson, Clement . . . 554

(See inside back cover)

Note to Librarians. The six bimonthly issues of "Applications and Industry," March 1961–January 1962, will also be available in a single volume (no. 80) entitled "AIEE Transactions—Part II. Applications and Industry," which includes all technical papers on that subject presented during 1961. Bibliographic references to Applications and Industry and to Part II of the Transactions are therefore equivalent.

Applications and Industry. Published bimonthly by the American Institute of Electrical Engineers, from 20th and Northampton Streets, Easton, Pa. AIEE Headquarters: 345 East 47th Street, New York 17, N. Y. Address changes must be received at AIEE Headquarters by the first of the month to be effective with the succeeding issue. Copies undelivered because of incorrect address cannot be replaced without charge. Editorial and Advertising offices: 345 East 47th Street, New York 17, N. Y. Nonmember subscription $8.00 per year (plus 75 cents extra for foreign postage payable in advance in New York exchange). Member subscriptions: one subscription at $5.00 per year to any one of three divisional publications: Communication and Electronics, Applications and Industry, or Power Apparatus and Systems: additional annual subscriptions $8.00 each. Single copies when available $1.50 each. Second-class mail privileges authorized at Easton, Pa. This publication is authorized to be mailed at the special rates of postage prescribed by Section 132.122.

Printed in United States of America

Number of copies of this issue 5,100

The Effect of Variable Plasma Conductivity on MHD Energy Converter Performance

W. B. COE
NONMEMBER AIEE

C. L. EISEN
NONMEMBER AIEE

THE CONCEPT OF a magnetohydrodynamic (MHD) energy converter presents a means of converting some of the kinetic and thermal energy of a plasma directly into useful electrical work. The principle upon which its operation is based is that of passing a moving plasma through a transverse magnetic field, thereby inducing an electromotive force (emf) in the plasma. By completing the circuit through an external load, a current is then permitted to flow. The analysis of such a device is complicated by the fact that the working fluid is a compressible medium, subject not only to the laws of compressible fluid flow, but to the laws of electromagnetism as well.

The potentialities of the MHD energy converter were investigated by Neuringer,[1] who applied the techniques of the calculus of variations to obtain an appropriate set of differential equations governing the conditions under which the maximum amount of useful electrical power could be generated. Although the equations he derived are quite general, they can be integrated in closed form only for special cases. One of the cases to which Neuringer applied these equations and obtained solutions was that of a converter with constant cross-sectional area and with constant plasma electrical conductivity. The results of this study showed that appreciable amounts of energy could be converted into useful electrical work.

A subsequent investigation undertook to obtain solutions to a set of differential equations governing the performance of the MHD converter. The equations were written for a varying channel cross-sectional area and allowed the electrical conductivity of the plasma to depend on its local thermodynamic properties. These equations were programmed for numerical solution on an International Business Machine Corporation (IBM) *704* digital computer. The results of the computation permit the investigation of the performance of the converter for other than maximum energy conversion. A preliminary study[2] confined itself to a constant cross-sectional area converter with an assumed constant plasma conductivity.

This paper reports the results of a similar performance study of a constant area channel in which the electrical conductivity of the plasma is dependent on its local thermodynamic properties.

Assumptions

There are a number of assumptions and approximations in the derivation of the equations governing the performance of the MHD energy converter. These include:

1. One-dimensional flow, i.e., the fluid dynamical state variables vary in the flow direction only, and not over the cross section.

2. Small magnetic Reynold's number, i.e., any effects on the fluid flow of secondary magnetic fields resulting from the induced current distribution are negligible, either because the secondary magnetic fields are small, or in the wrong direction to produce appreciable effects.

3. Negligible Hall current, i.e., the induced currents in the flow direction are small compared to those induced in the direction perpendicular to both the flow and to the magnetic fields.

4. The plasma behaves like a perfect gas.

5. The specific heats are constant.

6. The plasma is inviscid and non-heat-conducting so that the only dissipation is electrical.

7. The plasma is electrically neutral, i.e., no space charge sheaths are developed near the conducting walls.

8. Shock-free supersonic flow.

9. The channel walls are perfect electrical conductors.

Equations

Let the flow of the plasma through the channel be in the x direction, as shown in the schematic diagram of the converter, Fig. 1. The channel cross-sectional area is given by the product $y(x)z(x)$, where $y(x)$ is the channel height (or electrode spacing) and $z(x)$ is the channel depth, both being prescribed functions of x, subject to the conditions that dy/dx and dz/dx are everywhere sufficiently small as required by the one-dimensional flow approximation. The applied magnetic field, $B(x)$, is in the z direction and is also a prescribed function of x. (Although the results discussed in the next section are for constant channel cross-section and constant magnetic-field intensity, both will here be considered as varying functions.) With the given directions for the magnetic field and the flow, an electric field is induced in the y direction, so that a potential difference thereby appears across the electrodes. The electrodes then are the terminals of the converter, and a load

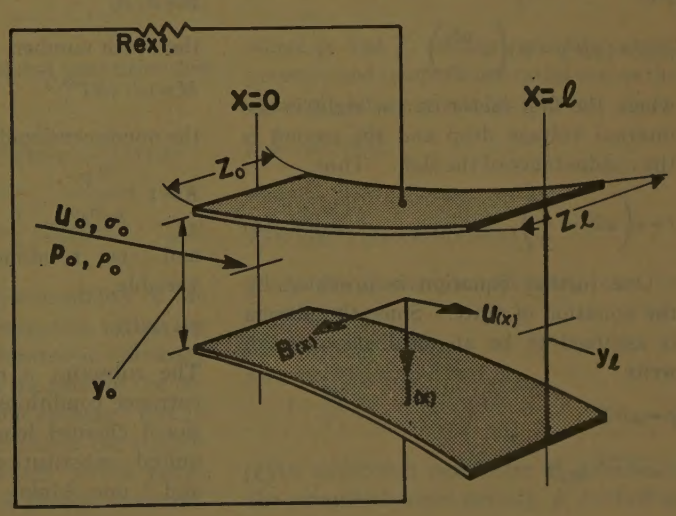


Fig. 1. Schematic of energy converter

Paper 60-1028, recommended by the AIEE Aerospace Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Pacific General Meeting, San Diego, Calif., August 8–12, 1960. Manuscript submitted May 12, 1960; made available for printing April 17, 1961.

W. B. COE and C. L. EISEN are with the Republic Aviation Corporation, Farmingdale, N.Y.

This work was sponsored by the United States Air Force Office of Scientific Research under contract no. *AF 49(638)-552* jointly with the Republic Aviation Corporation.

resistor connected across them will permit a current to flow.

The flow equations which apply to the plasma are the usual one-dimensional steady-state fluid dynamic equations of motion, with additional terms to account for the effect of the magnetic field on the flow. These equations are the continuity equation,

$$\frac{d}{dx}(\rho u y z)=0 \tag{1}$$

the momentum equation,

$$\rho u\frac{du}{dx}+\frac{dp}{dx}+jB=0 \tag{2}$$

and the energy equation,

$$\frac{d}{dx}\left[\rho u y z\left(c_p T+\frac{1}{2}u^2\right)\right]+jBuyz-\frac{j^2 yz}{\sigma}=0, \tag{3}$$

where ρ is the fluid density, u the fluid velocity, p the static pressure, j the induced current density, c_p the constant pressure specific heat, T the static temperature, and σ the local electrical conductivity of the plasma. In the momentum equation, the term jB is the reactive force per unit volume exerted by the magnetic field on the plasma. In the energy equation, the term $jBuyz$ is the rate at which work is done per unit length of channel by the plasma against the reactive force, and would be the power delivered to the external load in the absence of internal losses. The term j^2yz/σ is the power per unit length of channel dissipated within the plasma due to Joulean heating. For infinite conductivity, this term vanishes.

In addition to the fluid dynamic equations, an expression for the current density can be written for an assumed load on the converter. Let the load be such that the terminal voltage is V. Then the current in the slab of plasma of thickness dx, height y, and depth z is

$$jzdx=(uB\,y-V)\left(\sigma\frac{zdx}{y}\right)$$

where the first factor on the right is the internal voltage drop and the second is the conductance of the slab. Thus

$$i=\sigma\left(uB-\frac{V}{y}\right) \tag{4}$$

One further equation is provided by the equation of state. Since the plasma is assumed to be an ideal gas, one can write

$$p=\rho RT$$

$$=\frac{\gamma-1}{\gamma}\rho c_p T \tag{5}$$

where R is the gas constant for the plasma, and γ is the specific heat ratio.

Equations 1–5 constitute five equations in the five unknown functions ρ, u, p, T, and j. An immediate reduction to two equations in p and u can be made by using equations 4 and 5 and the integral of equation 1 to eliminate ρ, T, and j in equations 2 and 3. One obtains then

$$\frac{du}{dx}+\frac{yz}{m}\frac{dp}{dx}+\frac{B\sigma yz}{m}\left(uB-\frac{V}{y}\right)=0 \tag{6}$$

$$\frac{d}{dx}\left(\frac{\gamma}{\gamma-1}yzup+\frac{1}{2}mu^2\right)+V\sigma z\left(uB-\frac{V}{y}\right)=0 \tag{7}$$

where m, the mass flow rate, is the constant of integration of equation 1.

In order to make the parameteric studies discussed later, it is convenient to nondimensionalize equations 6 and 7. For this purpose the following nondimensional quantities are introduced:

the nondimensional channel length,

$$\delta=\frac{B_o^2 y_o z_o \sigma_o}{m}l \tag{8}$$

the nondimensional velocity,

$$U=u/u_o$$

the nondimensional static pressure,

$$P=p/\rho_o u_o^2$$

the nondimensional channel height,

$$Y=y/y_o$$

the nondimensional channel depth,

$$Z=z/z_o$$

the nondimensional electrical conductivity,

$$\bar{\sigma}=\sigma/\sigma_o \tag{9}$$

the nondimensional magnetic field strength,

$$\bar{B}=B/B_o$$

the Mach number,

$$M=u/(\gamma RT)^{1/2}$$

the nondimensional load parameter,

$$K=1-\frac{V}{u_o B_o y_o}$$

and the nondimensional independent variable,

$$\xi=(\delta/l)x$$

The subscript o refers to the channel entrance conditions, and l is the dimensional channel length. Making the required substitutions into equations 6 and 7, one obtains

$$\frac{dU}{d\xi}+YZ\frac{dP}{d\xi}+\bar{B}\bar{\sigma}YZ\left[\bar{B}U-\frac{1}{Y}(1-K)\right]=0 \tag{1}$$

and

$$\frac{d}{d\xi}\left(\frac{\gamma}{\gamma-1}YZUP+\frac{1}{2}U^2\right)+(1-K)\bar{\sigma}Z\left[\bar{B}U-\frac{1}{Y}(1-K)\right]=0 \tag{1}$$

In equations 10 and 11, Y and Z a prescribed functions of ξ, $\bar{\sigma}$ is considere a known function of Y, Z, U, and and K is the load parameter which ca assume prescribed values between ze (open external circuit) and unity (shor circuited external circuit). The dete mination of $U(\xi)$ and $P(\xi)$ over the inte val $0\leq\xi\leq\delta$ constitutes the solution equations 10 and 11.

To obtain an expression for the ele trical conductivity it is assumed that t entire current is carried by the electron Thus

$$\sigma=\frac{j}{E}=\frac{n_e e u_e}{E} \tag{1}$$

where n_e is the number density of ele trons, e the electronic charge, u_e the dr velocity of electrons through the ga and E the electric field intensity. general the electron mobility, u_e/E, limited by collisions with neutral g atoms and by long-range Coulomb inte actions with other charged particles. T former process predominates in a weak ionized gas, and the latter in a strong ionized gas. If $1/\sigma_1$ is the resistivity the gas due to electron collisions wi neutral gas atoms and $1/\sigma_2$ the resistivi due to long-range Coulomb interactions is assumed then, that for a gas of inte mediate ionization, in which both pro esses are significant, the resistivity given by the sum of the resistivities due the two processes:[3]

$$\frac{1}{\sigma}=\frac{1}{\sigma_1}+\frac{1}{\sigma_2} \tag{1}$$

For a weakly ionized gas in which t electric field is not large enough to cau ionizing collisions, the collision encount between electrons and neutral gas ato are mostly elastic. In such a gas Cha man's theory[4] for electron mobility giv

$$\frac{u_e}{E}=\frac{2}{3}\sqrt{\frac{2}{\pi}}\frac{e}{nQ}\left(\frac{1}{m_e kT}\right)^{1/2}$$

$$=\frac{2}{3}\sqrt{\frac{2}{\pi}}\frac{e}{Qp}\left(\frac{kT}{m_e}\right)^{1/2} \tag{1}$$

where n is the number density of g atoms, Q the electron-gas atom collisi cross section, k Boltzmann's constant, the gas temperature, and p the pressu

Although Q is a function of temperature, or this study it is taken as a suitable onstant for the temperature ranges of nterest.)

The number density of electrons as required by equation 12 can be obtained rom Saha's equation,[5] assuming that the as as a whole is electrically neutral and hat the electrons are in equilibrium with he gas. It is further assumed that for emperatures of interest one can ignore all nternal states of the gas atoms and their ons except the states of lowest energy. aha's equation can then be written

$$n_s^2 = n\frac{g_e g_i}{g_a}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2}\exp(-\varphi/kT)$$

$$= p\frac{g_e g_i}{g_a}\left(\frac{2\pi m_e k^{1/3}T^{1/3}}{h^2}\right)^{3/2} exp(-\varphi/kT) \quad (15)$$

where $g_e(=2)$ is the multiplicity of the lectron, g_i and g_a are the respective multiplicities of the ground states of the ion nd neutral atom, h is Planck's constant, nd φ is the ionization potential.

Substituting into equation 12 for the nobility as given by equation 14, and for he electron particle density as given by quation 15, one has for a weakly ionized as

$$\sigma_i(p, T) = \frac{(2^9\pi)^{1/4}}{3}\frac{e^2 m_e^{1/4}k^{3/4}}{h^{3/2}}\times \left(\frac{g_e g_i}{g_a}\right)^{1/2}\frac{1}{Q}\frac{T^{3/4}}{p^{1/2}} exp(-\varphi/2kT) \quad (16)$$

For a fully ionized gas, the electrical onductivity, as given by Spitzer and Härm,[6] is, in mks (meter-kilogram-second) nits,

$$\sigma_2 = \frac{2^{12/2}\pi^{1/2}\kappa_0^2\gamma_E k^{3/2}}{m_e^{1/2}Ze^2}\cdot\frac{T^{3/2}}{\ln(qc^2)} \quad (17)$$

where κ_0 is the permittivity of free space, γ_E a correction factor to account for electron-electron encounters, and Ze the ionic harge:

$$q = \frac{4\pi\kappa_0^{3/4}m_e(kT)^{1/2}}{Z(Z+1)^{1/2}e^3 n_e^{1/2}} \quad (18)$$

nd

$$c = \left(\frac{3kT}{m_e}\right)^{1/2} \quad (19)$$

or a singly ionized gas, $Z=1$ and $\gamma_E=$ .5816.

For a fully ionized gas,

$$n_e = \frac{p}{2kT} \quad (20)$$

Equations 17–20 combine to express σ_2 s a function of p and T. The results together with equation 16 are substituted nto equation 13 to provide the conductivity function, which is then nondimensionalized according to equation 9, where σ_0 is the conductivity evaluated at the entrance to the channel. Equations 10 and 11, however, require that the conductivity be expressed as a function of the nondimensional variables. Thus the substitutions

$$p = \gamma M_0^2 p_0 P$$

and

$$T = \gamma M_0^2 T_0 YZUP$$

must be made in $\bar{\sigma}$.

The power output of the converter is given by

$$\hat{P} = V\int_0^l jzdx = V\int_0^l \sigma z\left(uB - \frac{V}{y}\right)dx$$

where the current density is obtained from equation 4. Substituting for the integrand from equation 7, the output power can be expressed in the following convenient form:

$$\hat{P} = h_T(0) - h_T(l)$$

where

$$h_T(x) = \frac{\gamma}{\gamma-1}yzup + \frac{1}{2}mu^2$$

is the stagnation enthalpy flux at x. A measure of the output power is therefore provided by the conversion effectiveness, η_c, which is defined as the ratio of the output power to the stagnation enthalpy flux at the entrance to the channel:

$$\eta_c = \frac{\hat{P}}{h_T(0)} = 1 - \frac{h_T{}^l}{h_T(0)}$$

$$= 1 - \frac{T_T{}^l}{T_T(0)} \quad (21)$$

where T_T is the stagnation temperature, c_P being assumed constant. The nondimensional stagnation enthalpy flux is defined by

$$H_T(\xi) = \frac{h_T(x)}{mu_0{}^2} \quad (22)$$

In terms of nondimensional quantities this can be written

$$H_T(\xi) = \frac{\gamma}{\gamma-1}Y(\xi)Z(\xi)U(\xi)P(\xi) + \frac{1}{2}[U(\xi)]^2$$

$$= \frac{1}{\gamma-1}\left[\frac{U(\xi)}{M(\xi)}\right]^2 + \frac{1}{2}[U(\xi)]^2 \quad (23)$$

A measure of the irreversibility of the flow due to Joulean dissipation within the fluid is given by the isentropic efficiency η_s, which is defined as the ratio of the decrease in the stagnation enthalpy flux in the channel to the decrease in the stagnation enthalpy flux which would result from an isentropic flow process terminating in an exit stagnation pressure equal to that of the actual flow. Thus

$$\eta_s = \frac{h_T(0) - h_T(l)}{h_T(0) - h'_T(l)} = \frac{\eta_c}{1 - \dfrac{h'_T(l)}{h_T(0)}} = \frac{\eta_c}{1 - \dfrac{T'_T(l)}{T_T(0)}}$$

again assuming the constancy of the specific heat. Here $h'_T(l)$ and $T'_T(l)$ are respectively the exit stagnation enthalpy flux and temperature for the isentropic flow process. One can also write for the isentropic process

$$\frac{T'_T(l)}{T_T(0)} = \left[\frac{p_T(l)}{p_T(0)}\right]^{(\gamma-1)/\gamma}$$

where p_T is the total pressure, and where $p_T(l)$ is the exit total pressure for both the actual flow and the isentropic flow. Thus

$$\eta_s = \frac{\eta_c}{1 - \left[\dfrac{p_T(l)}{p_T(0)}\right]^{(\gamma-1)/\gamma}}$$

To compute η_s, therefore, one must express the total pressure ratio, $p_T(l)/p_T(0)$, across the channel in terms of the solutions of equations 10 and 11. To do this one notes that the stagnation to static pressure and temperature are given respectively by

$$\frac{p_T}{p} = \left[1 + \frac{1}{2}(\gamma-1)M^2\right]^{\gamma/(\gamma-1)}$$

and

$$\frac{T_T}{T} = 1 + \frac{1}{2}(\gamma-1)M^2 \quad (24)$$

so that

$$\frac{p_T(l)/p_T(0)}{p(l)/p(0)} = \left[\frac{T_T(l)/T_T(0)}{T(l)/T(0)}\right]^{\gamma/(\gamma-1)}$$

or

$$\left[\frac{p_T(l)}{p_T(0)}\right]^{(\gamma-1)/\gamma} = \frac{p_r{}^{(\gamma-1)/\gamma}}{T_r}\times\frac{T_T(l)}{T_T(0)}$$

$$= \frac{p_r{}^{(\gamma-1)/\gamma}}{T_r}(1-\eta_c)$$

where p_r and T_r are the respective static pressure and temperature ratios across the channel and are given by

$$p_r = \gamma M_0^2 P(\delta)$$

and

$$T_r = \gamma M_0^2 Y(\delta)Z(\delta)U(\delta)P(\delta)$$

Thus the isentropic efficiency can be written

$$\eta_s = \frac{\eta_c}{1 - \dfrac{p_r{}^{(\gamma-1)/\gamma}}{T_r}(1-\eta_c)}$$

An additional parameter of interest is the generated power density, $\hat{p}$, defined as

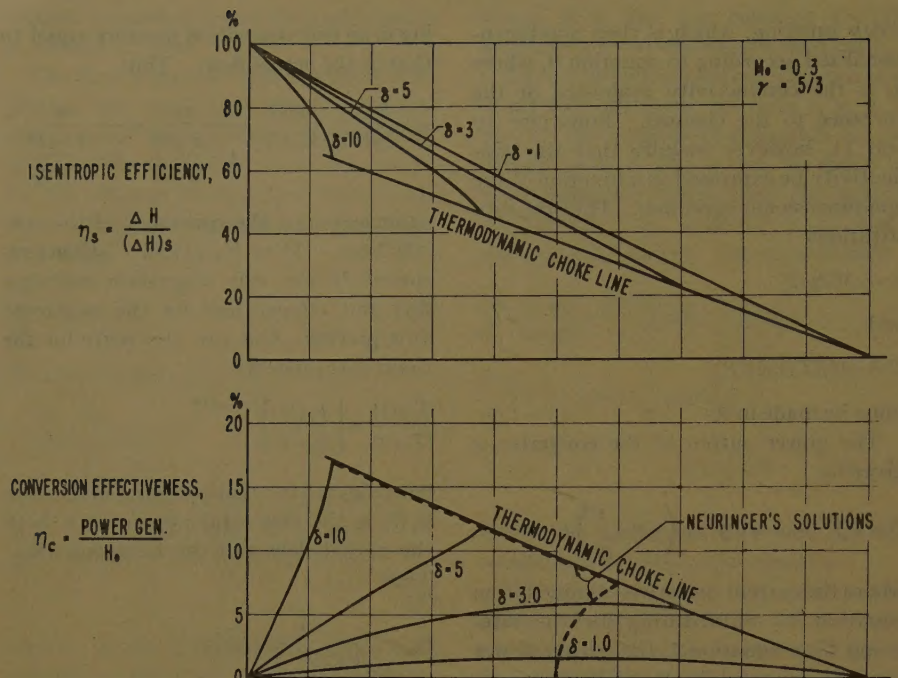


Fig. 2. Subsonic performance, constant area, constant conductivity

the ratio of the power output to the volume, v, of the channel:

$$\hat{p} = \frac{h_T(0)\eta_c}{v} \tag{25}$$

From equations 22 and 23, it is seen that the entrance enthalpy flux can be written

$$h_T(0) = mu_o^2\left[\frac{1}{(\gamma-1)M_o^2} + \frac{1}{2}\right] \tag{26}$$

The initial velocity may be expressed as

$$u_o = (\gamma RT_o)^{1/2} M_o \tag{27}$$

and by equation 8, the mass flow ratio may be given by

$$m = \frac{B_o^2 y_o z_o \sigma_o l}{\delta} \tag{28}$$

For the volume of the channel, one has

$$v = \int_0^l yz\,dx$$

$$= \frac{y_o z_o l}{\delta}\int_0^\delta YZ\,d\xi$$

$$= \frac{y_o z_o l}{\delta}\bar{v} \tag{29}$$

where a nondimensional volume is defined by

$$\bar{v} = \int_0^\delta YZ\,d\xi \tag{30}$$

Substituting the expressions given by equations 26–29 into equation 25, one obtains

$$\hat{p} = \frac{\gamma}{\gamma-1} B_o^2 \sigma_o RT_o\left[1 + \frac{1}{2}(\gamma-1)M_o^2\right]\frac{\eta_c}{\bar{v}}$$

$$= \frac{\gamma}{\gamma-1} B_o^2 \sigma_o RT_T(0)\frac{\eta_c}{\bar{v}} \tag{31}$$

Except for the channel shape, the magnetic field distribution, and the gas dat (γ, R, g_i, g_a, φ, and Q), the input param eters needed to compute η_c and η_s ar M_o, δ, K, T_o, [or $T_T(0)$], and p_o [or $p_T(0)$ the temperature and pressure bein needed only in the conductivity functio The computation of $\hat{p}$ also requires B_o.

Discussion and Results

The nondimensional equations 10 an 11 were solved numerically, using the IBM *704* digital computer. The performanc of the MHD energy converter was ca culated for constant channel cross-sec tional area ($Y=1$ and $Z=1$) and for constant magnetic field ($\bar{B}=1$). Th specific heat ratio γ was taken as 5/3 representing either a monatomic gas or fully ionized plasma.

The performance was computed for constant area energy converter; first wit an assumed constant plasma conductivit ($\bar{\sigma}=1$), then, for a specified working flui with the plasma conductivity, given b equations 9, 13, 16, and 17 as a functio of the local thermodynamic properties o the plasma. The performance was ob tained for a variety of subsonic and supe sonic entrance Mach numbers M_o, an for a variety of nondimensional chann lengths δ, the load parameter K, rangin from open circuit ($K=0$) to short circu ($K=1$).

Figs. 2 and 3 present the performanc

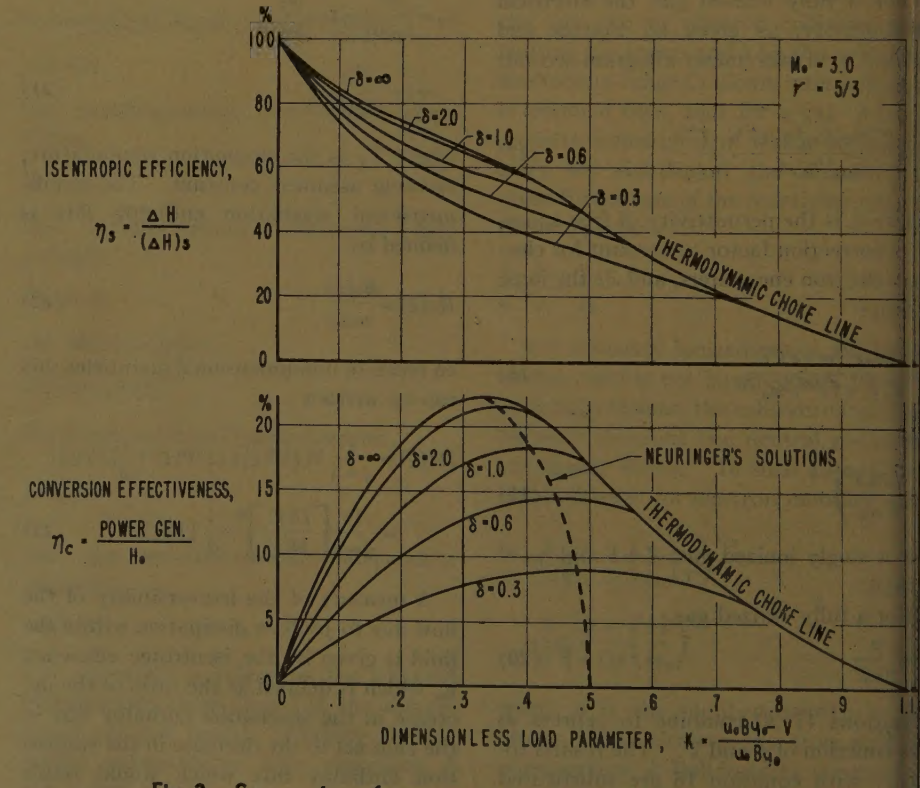


Fig. 3. Supersonic performance, constant area, constant conductivity

the constant area energy converter for assumed constant plasma conductivity.

The performance of a subsonic energy nverter is shown in Fig. 2 for an entnce Mach number of 0.3. The isenpic efficiency η_s and the conversion ectiveness η_c are plotted as functions of load parameter K for a range of nnel lengths δ. For small values of the conversion effectiveness increases d the isentropic efficiency decreases as is increased from zero. At approxiately $K=1/2$ the conversion effectiveess is a maximum, and further increases K cause η_c to decrease until at $K=1$ energy converted becomes zero. The ntropic efficiency continually decreases th increasing K until it also goes to zero $K=1$. When $K=1$, the converter is ort-circuited and all the gas dynamic ergy removed from the plasma is rned in the form of Joulean dissipation. r larger values of the channel length, δ, load parameter, K, cannot be inased beyond a certain limiting valve. is limit is imposed by the flow becomg sonic at the exit of the channel. hen this occurs the flow is choked. ie locus of this condition is indicated by e line labeled "thermodynamic choke e." It is thus seen that for shorter annels the maximum energy conversion curs when the loading is chosen in a anner which best matches the energy tput to the Joulean dissipation; for nger channels the maximum energy nversion is dictated by the thermonamic choking of the flow.

At a fixed value of loading, increasing e length of the converter causes an inease in η_c and a decrease in η_s. It also creases the exit flow Mach number until e choke condition is reached. For the bsonic converter the flow velocity ineases along the channel length. This turn produces a larger induced emf and quires, therefore, larger internal voltge drops to match the terminal voltage ecified by K. Each subsequent inement in length added to the channel nverts additional energy, but does so th poorer local efficiency.

The performance for the supersonic nverter is shown in Fig. 3 for an enance Mach number of 3.0. The hermodynamic choke line" again notes the operating conditions where e exit flow Mach number is unity. As posed to the subsonic converter, at a ed value of loading, both the conver-on effectiveness and the isentropic iciency of the supersonic converter inease as the length of the channel is ineased. For this case the flow velocity decreases along the length of the channel, causing a reduction in the induced emf. This requires, therefore, a smaller internal voltage drop in order to match the specified terminal voltage. Thus subsequent increments of length added to the channel convert energy with a local efficiency that is better than the average for the channel. For lightly loaded operation, i.e., small values of K, the drop in velocity causes the induced emf to approach the terminal voltage asymptotically. When this occurs further lengthening of the channel produces no change in the flow conditions or the energy converted. The line denoted by $\delta=\infty$ indicates this condition.

Figs. 2 and 3 also show that the maxima of the conversion effectiveness agree with the results of Neuringer's optimum power analysis.[1] An interesting characteristic is the value of K at which these maxima occur. If the plasma acted as a perfectly incompressible fluid, or a solid conductor, maximum conversion would occur when the energy delivered to the external load is equal to that lost by Joulean dissipation within the converter. For an incompressible fluid this condition is obtained when $K=1/2$, since the induced emf is constant down the channel. As seen in Figs. 2 and 3, for short channels both the subsonic and supersonic converter also require $K\sim1/2$ for maximum conversion. As the channel length of both converters is increased the value of K at maximum conversion deviates from $K=1/2$, the subsonic converter requiring larger values of K in the unchoked region, and the supersonic converter always requiring smaller values of K. For these larger channel lengths the compressibility effects are significant and the local loading of each element of the channel is related to that of all the other elements through the gas dynamics involved. As a result, the condition for maximum conversion is no longer simply that required by a solid or incompressible conductor.

In order to study the effect of a variable conductivity plasma upon the performance of the MHD energy converter, a choice must be made of the working fluid and its initial conditions. For this investigation cesium vapor was chosen as the working fluid, the choice being dictated by its low value of ionization potential, $\varphi=3.87$ volts, which will provide reasonable values of plasma conductivity for temperatures in the thermal range. The performance was computed for two values of entrance stagnation temperature, T_{T0}. The lower value, $T_{T0}=2{,}000$ K (degrees Kelvin), represents the lowest usable plasma temperature at which present-day materials can conceivably operate. The higher value of 5,000 K has been included to indicate the increased performance capability that could be obtained, without regard for the materials

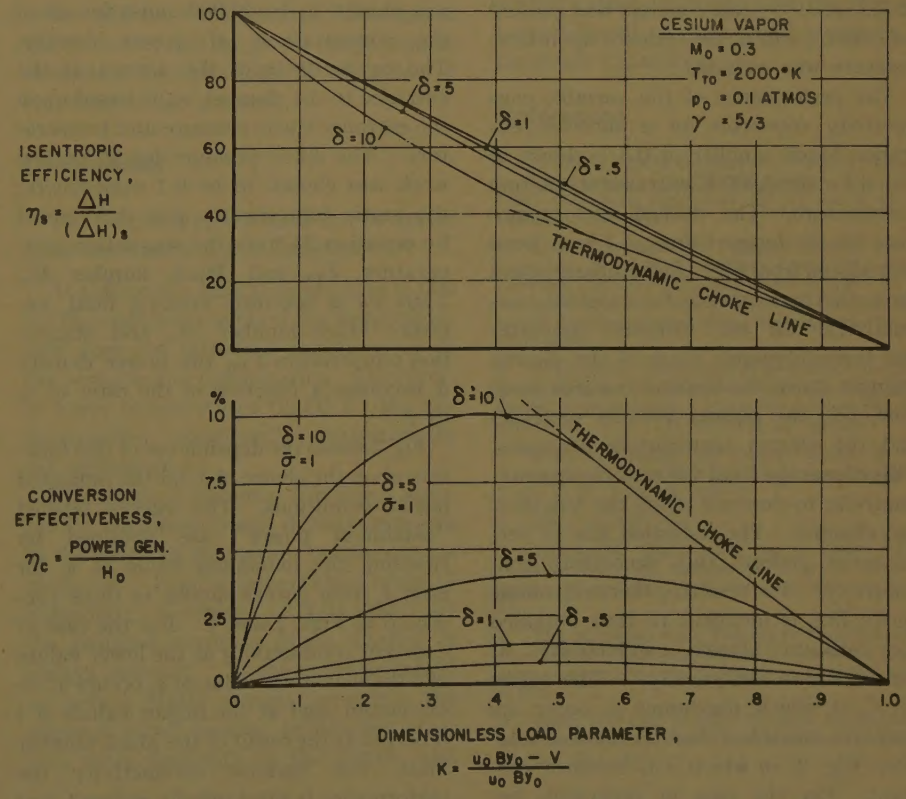


Fig. 4. Subsonic performance, constant area, variable conductivity

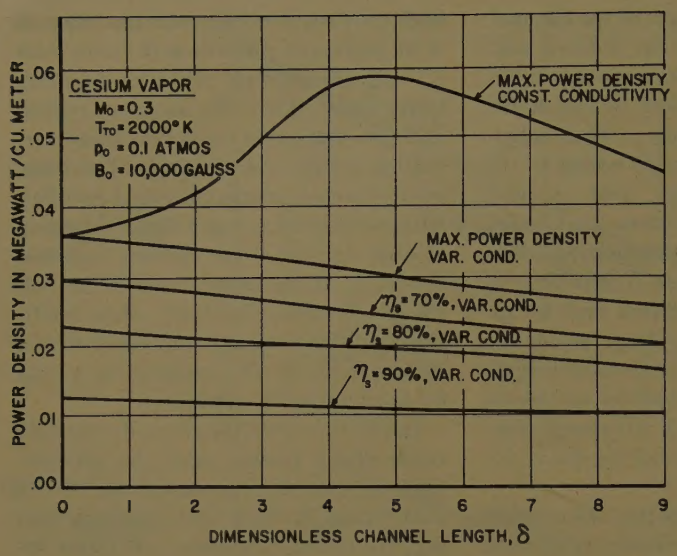


Fig. 5. Power density as a function of channel length

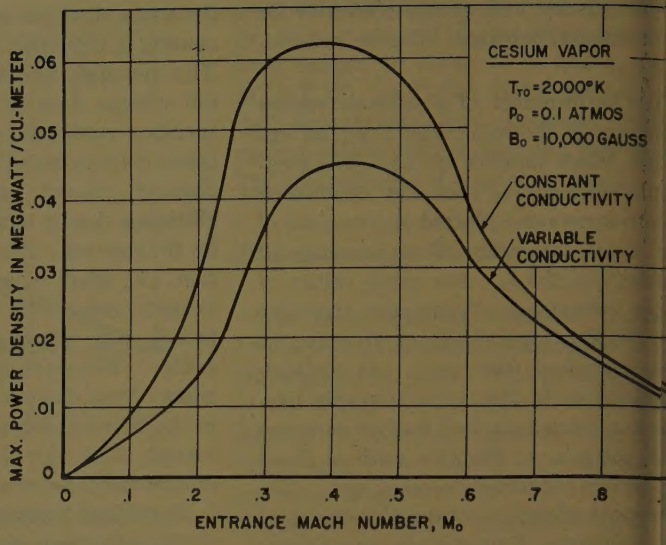


Fig. 6. Maximum power density for T_{To} = 2,000 K

problem. All the performance computations reported here, for variable plasma conductivity, have been made for an entrance static pressure, p_o, of 0.1 atmosphere. This represents the lowest value for the present choice of working fluid and and temperatures that will not violate the assumption of a scalar plasma conductivity.[7]

The conductivity function requires the collision cross section of electrons with atoms as a function of temperature. Because good data were not available in the required temperature range, a value of 3.25×10^{-18} square meters was chosen[8] for cesium, and no dependence upon temperature was included.

The performance of the variable conductivity converter for a subsonic entrance Mach number of 0.3 is shown in Fig. 4 for the 2,000 K entrance stagnation temperature. The dashed performance lines for the larger values of δ have been reproduced from Fig. 2 for comparison and show the performance for constant conductivity. In the subsonic converter the thermodynamic state of the plasma changes down the channel towards sonic flow, i.e., the plasma velocity increases and the plasma temperature decreases. This causes the local value of plasma conductivity to decrease along the length of the channel. The indicated loss in performance reflects this decreasing conductivity. The resulting thermodynamic choke line is identical to that obtained for constant plasma conductivity, as can be shown analytically. The values of K at which maximum η_c occur are lower for variable σ than for the computation (Fig. 2) in which σ is assumed constant. For the case of supersonic entrance Mach numbers, the local value of conductivity increases along the length of the channel and results in an increase in converter performance. The consequences of this characteristic will be indicated in some of the subsequent results.

For a constant area channel the nondimensional volume, as given by equation 30, reduces to $\bar{v}=\delta$. The power density $\hat{p}$ given by equation 31 then becomes

$$\hat{p}=\frac{\gamma}{\gamma-1}B_o^2\sigma_o RT_{To}\frac{\eta_c}{\delta} \qquad (32)$$

The constant magnetic field strength, B_o, was chosen to be 10,000 gauss for all of the computations of power density. The conductivity of the plasma at the entrance to the channel, σ_o, is based upon the entrance static pressure and temperature. The static pressure for all of this work was chosen to be 0.1 atmosphere; the static temperature was determined by equation 24 from the stagnation temperature T_{To} and Mach number M_o. Thus for a specified working fluid, entrance Mach number M_o and stagnation temperature T_{To}, the power density $\hat{p}$ becomes a function of the ratio of η_c to δ.

Fig. 5 shows the dependence of this function upon the choice of δ, for the indicated initial conditions. The curves labeled "maximum power" are obtained by choosing the maximum value of η_c for each δ, from curves similar to those presented in Figs. 2 and 4. For the case of constant conductivity at the lower values of δ the maximum value of η_c occurs as an extremum, and at the higher values of δ it occurs as the result of the Mach number limit. For variable conductivity the performance is substantially reduced, and at all δ's calculated, η_c occurs as an extremum. Also included in Fig. 5 a power densities for various consta values of isentropic efficiency, η_s. Th curves are included to show how t power density would be decreased from maximum if high isentropic efficie were an important design consideration

For temperatures of the order of 2,000 the conductivity of a cesium plasma v vary predominantly as an exponen function of temperature as indicated equation 16. The entrance static te perature decreases for increasing values entrance Mach numbers and fixed val of T_{To}. Although this results in a la reduction in σ_o, the maximum power d sity for small subsonic Mach numbers, creases with Mach number, as shown Fig. 6 for T_{To}=2,000 K. As seen fr equation 32 this is possible only beca the ratio of η_c to δ increases faster than decreases. As the entrance Mach nu ber is further increased a point is reach after which the maximum power dens decreases, since at the high Mach nu ber, the decrease in σ_o is the dominant f tor in equation 32. The value of po density for supersonic entrance M numbers has not been calculated for T_T 2,000 K due to the extremely low value plasma conductivity. The sensitive ture of the conductivity function at t peratures in this range causes an app ciable variation in the power density t can be achieved when the plasma c ductivity is allowed to vary in a reali manner as opposed to holding it a const The performance shown in Fig. 6 dicates an overestimation of about 4 in power density if the conductivity assumed to be constant.

The performance shown in Fig. 7 is an entrance stagnation temperature

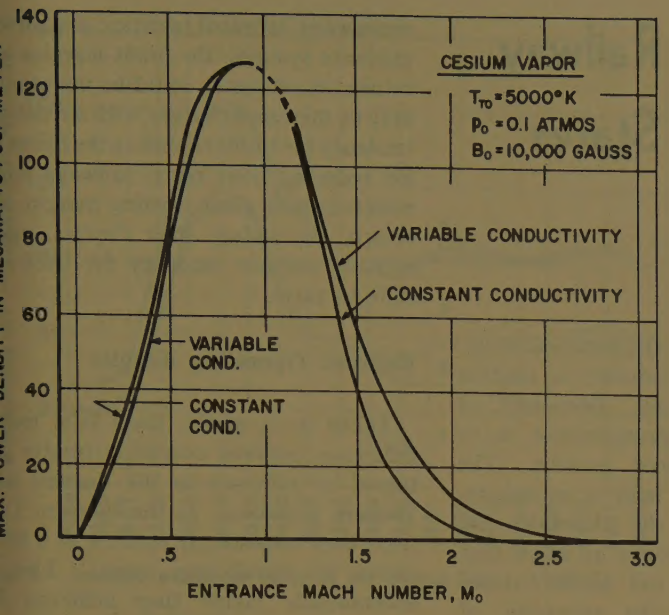


Fig. 7. Maximum power density for T_{To} = 5,000 K

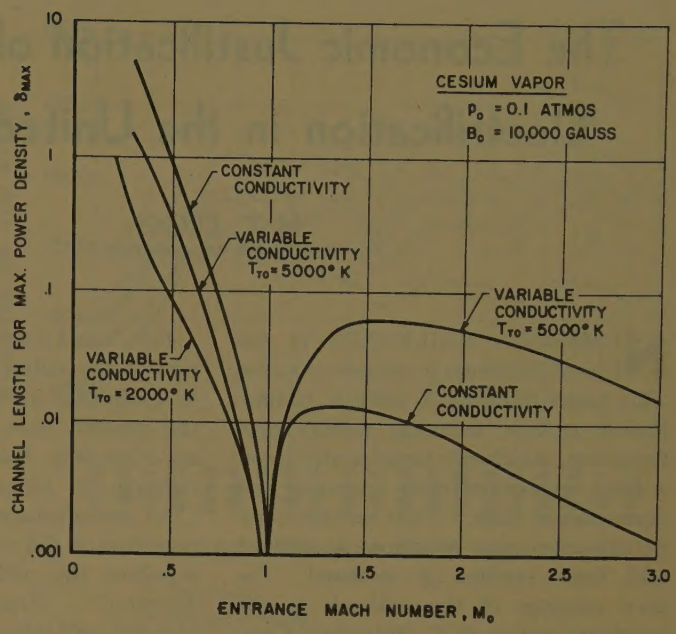


Fig. 8. Converter length for maximum power density

,000 K. The large increase in the level f power density is mainly due to the exonential increase in the level of conducivity, although the increased stagnation emperature itself causes an increase by a actor of 2.5. This indicates that for mall increases in the stagnation temperaure above 2,000 K large gains in power ensity could be achieved. For a stagnaion temperature of 5,000 K the variaion of σ with changes in temperature s much smaller than at the lower temperaure. This is indicated in Fig. 7 by the mall difference between the performance ith and without variable conductivity. lso because of this smaller variation in the performance computation may be arried into the supersonic range. The aximum value of power density now ccurs at a higher value of entrance Mach umber.

The length of channel, δ, required o achieve the maximum power densities hown in Figs. 6 and 7 is plotted in Fig. 8 s a function of entrance Mach number. or constant plasma conductivity the equired δ is independent of temperature. he effect of variable conductivity is to horten the channel length for subsonic Iach number and lengthen the channel or supersonic Mach number.

With a specified value of 2,000 K for the ntrance stagnation temperature and an ntrance static pressure of 0.1 atmoshere the maximum power density, for a ariable plasma conductivity, occurs or an entrance Mach number of $M_o \sim$ .45, as shown in Fig. 6. The nondimenonal channel length required for maximum power density is $\delta \sim 0.1$, as given by the length functions of Fig. 8. From the definition of δ (equation 8) one can write

$$l = \frac{\rho_o u_o \delta}{B_o^2 \sigma_o} = \frac{M_o \delta p_o \gamma^{1/2}}{\sigma_o B_o^2 (RT_o)^{1/2}} \qquad (33)$$

where l is the dimensional length of the channel. Substituting into equation 33 for the case cited above, one obtains $l \sim 50$ centimeters. For a channel height and width equal to its length the output power will be $\sim$5 kilowatts. The size of the converter and the power output can be controlled by the choice of the value of entrance static pressure.

Conclusions

The following conclusions can be drawn from the results of this analysis:

At sufficiently high temperatures (5,000 K) the constant conductivity approximation does not introduce serious error in the performance computation. However, at lower temperatures (2,000 K) the approximation significantly overestimates the performance.

The required entrance flow Mach number for maximum power density increases as the entrance stagnation temperature is increased. For the temperature range of immediate interest (2,000–5,000 K) the required entrance Mach number is always subsonic. In the construction of an actual MHD energy converter, the subsonic Mach number requirement permits more flexibility in the design, and lends more validity to the approximations which are made in the analysis than would be true for a supersonic energy converter.

The attainable power density increases sharply with increases in entrance stagnation temperature. Since the maximum temperature that an MHD energy converter can operate at is limited by the capabilities of the materials involved, the reward of large increases in performance for small increases in operating temperatures provides a strong incentive for sophistication in design and for the accomplishment of even small increases in materials capability.

References

1. Optimum Power Generation Using a Plasma as a Working Fluid, J. L. Neuringer. *Journal of Fluid Mechanics*, New York, N. Y., vol. 7, pt. 2, 1960, pp. 287–301.

2. The Performance Characteristics of a Constant Area MHD Energy Converter with Assumed Constant Plasma Conductivity, C. L. Eisen, W. B. Coe. *Report no. PPL-TN-60-7*, Republic Aviation Corporation, Farmingdale, N. Y., Mar. 1960.

3. Electrical Conductivity of Highly Ionized Argon Produced by Shock Waves, S. Lin, E. L. Resler, A. Kantrowitz. *Journal of Applied Physics*, New York, N. Y., vol. 26, 1955, pp. 95–109.

4. Kinetic Theory of Gases (book), E. H. Kennard. McGraw-Hill Book Company, Inc., New York, N. Y., 1938, pp. 470–73.

5. *Ibid.*, pp. 243–45.

6. Transport Phenomena in a Completely Ionized Gas, L. Spitzer, R. Härm. *Physical Review*, New York, N. Y., vol. 89, 1953, pp. 977–81.

7. The Mathematical Theory of Non-Uniform Gases (book), S. Chapman, T. G. Cowling. Cambridge University Press, Cambridge, England, 1952, p. 328.

8. Basic Data of Plasma Physics (book), Sanborn C. Brown. John Wiley & Sons, Inc., New York, N. Y., 1959, p. 6.

The Economic Justification of Railway Electrification in the United States

H. C. CROSS
ASSOCIATE MEMBER AIEE

NUMEROUS RAILROADS in the United States are in economc distress. This paper presents a modern technological package, including railway electrification, which will permanently lower a large segment of their operating and maintenance costs. Their current competitive economic situation is outlined and their position is reviewed. The poor earnings of the railroads in the northeast sector—ICC (Interstate Commerce Commission) Eastern District—relative to those in the other sectors—ICC Western District—are shown.

New types of communication and signal-interlocker systems, electric locomotives converted from existing diesel-electric locomotives, and the commercial-frequency high-voltage system of electrification are included in the technological package, which is described in detail, together with its advantages.

An economic study shows the annual economic results for one year before conversion and each of 12 years thereafter, using a typical Eastern District railroad, operating inland from the eastern seaboard, as its basis. For purposes of comparison, results are shown with both dieselization and electrification. The package with electrification produces major and permanent savings not offered with dieselization. Moreover, the net capital investment is less with electrification; major increases are made in freight train speeds, and equal or increased capacity is provided.

For railroads in the Eastern District which now have relatively heavy traffic, and for the heavy traffic routes which will develop from future railway mergers, this package is of great economic importance.

Railroads as a Growth Industry

The railroad transportation industry in the United States is regarded as unprogressive. Freight traffic, as measured in ton-miles, remains more or less unchanged, but passenger traffic, measured in terms of passenger-miles, is continuously decreasing, even though, in over-all terms, both intercity freight and passenger traffic is steadily increasing.

Figs. 1 and 2 show the historical situation of the railway industry in relation to competing industries. Obviously, at the present time, transportation is in an effectively balanced position. The railways no longer hold a monopoly. "The manufacturer, the wholesaler, the retailer, and the customer all know that somehow the milk will arrive on the doorstep." From the viewpoint of national self-interest, the railway industry should share in the increasing traffic by, at least, maintaining its existing percentages of the total.

Fig. 3 shows the operating results for Class I railroads and the competing industries. The railways' profit margin and return on net assets have declined steadily since 1955. The 1959 margin was the same as for total manufacturing, twice that for motor carriers and air transport, and only slightly lower than that for interstate bus companies and inland water carriers. The return on net assets of competing companies ranges from 1.8 to 5.1 times that of the railways. This is basically the result of (1) overstatement of the railroad industry net worth growing out of conservative depreciation policies and (2) the capital investment in right of way, tracks, and signals, for which competing industries make no comparable investment. This situation is clearly brought out in the graph at the right-hand side of Fig. 3, showing net assets per dollar of sales. Trucks, buses, and airplanes operate at abnormally low net assets per dollar of sales because of indirect subsidies which, as national policy, were granted them in order to break the previously existing railway transportation monopoly.

If the competing industries were to pay the full costs of independent throughways, waterways, terminal facilities, and airw guidance systems, the profit margins a return on net assets could be maintain only by increased charges, with a result tendency for traffic to shift to the railwa By reducing their costs, railways cou maintain both their existing margin a return, and reduce their charges, aga with a resultant tendency for traffic shift to them.

Railroad Operating Results

Taken as a whole, there is a ma difference between operating results o tained by railroads in the Eastern a Western Districts. In the Western D trict, they consistently operate at a le above the break-even point. This substantially higher than achieved Eastern District railways, even thou their traffic in gross ton-miles per mile route is about one third higher and th total revenue per mile of route is appro mately 80% higher than in the West District. Western's margin on sales d ing 1958 was about 8.3% as compa with Eastern's margin of approximat 2.5%.

Relatively, Eastern railways have grea trackage per mile of route, particularly to other main tracks and way- and ya switching tracks. The particularly hi proportion of yard-switching tracks dicates relatively excessive terminal co undoubtedly arising from duplication facilities.

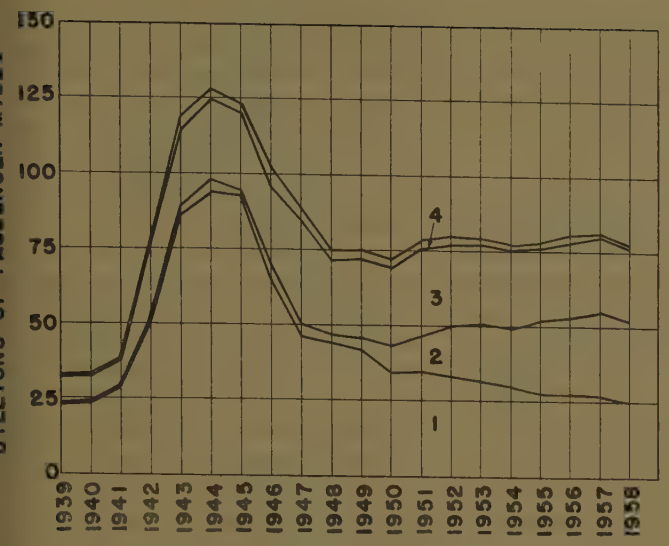
They encounter more severe compe tion from (1) trucks operating over hi ways provided at public expense and shipping on inland waterways and St. Lawrence Seaway route into the mid west. As a whole, they normally oper uncomfortably close to the break-e position as to earnings. During a recession or even temporary major s pension such as was caused by the st strike in 1959, individual roads imm diately take a loss. In view of the margin on sales, rate reductions to hi existing traffic must be limited. On contributed margin—or out-of-pocke basis, rate reductions must either attr additional customers without disturb the existing volume, or these reducti be in such large volume as will m than offset any reduction of income existing traffic.

The imperative need seems to be red tion in operating and maintenance co attained with a minimum of capital penditure. Also, to assist in attract additional traffic to the railroads, grea increased speed is required of frei trains.

Paper 61-208, recommended by the AIEE Land Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 29–February 3, 1961. Manuscript submitted November 23, 1960; made available for printing March 14, 1961.

H. C. Cross is with Westinghouse Electric International Company, A Division of Westinghouse Electric Corporation, New York, N. Y.

The author acknowledges assistance rendered in various ways by H. F. Brown, L. W. Birch, T. A. Benner, E. K. Bloss, R. L. Kimball, C. R. Kingston, A. G. Oehler, and J. R. Shepard.



g. 1. Intercity passenger traffic in the United States, 1939–58

1—Railroads
2—Airplanes
3—Buses
4—Water carriers and electric railways
(Numbers refer to area between lines)

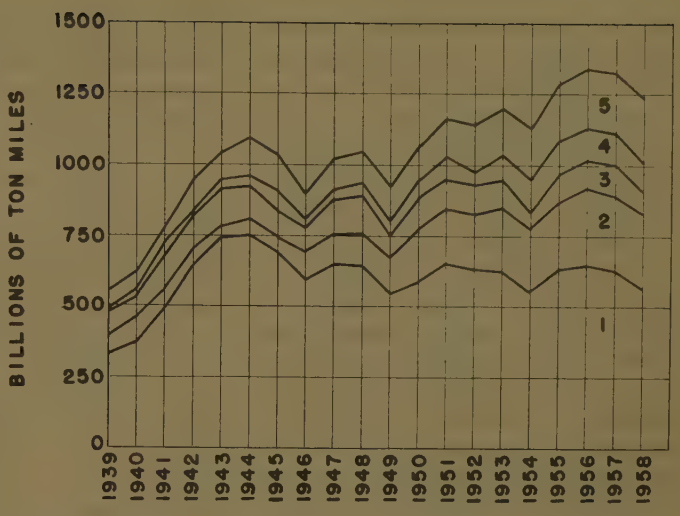


Fig. 2. Intercity freight traffic in the United States, 1939–58

1—Railroads
2—Motor carriers
3—Great Lakes
4—Rivers and canals
5—Oil pipelines
(Numbers refer to areas between lines)

ne Technological Package

Suitable Standard System for Railway Electrification

A standard electrification system hav- g a world-wide application has come to existence in recent years, making use a commercial-frequency single-phase ,000-volt power supply to a catenary- spended trolley over railway tracks. his system became practical with the velopment of rectifier-type locomotives which the 25,000-volt power supply is epped down to a voltage suitable for plication to rectifiers supplying a lsating direct current to standard series mmutator traction motors.

The ignitron rectifier locomotive ini- lly was developed in the United States d was adopted extensively by French anufacturers, both for use in their own untry and for export. The British ilways also have adopted this system electrification of their trackage, and actively engaged in installing the wer supply system in various regions. eir national manufacturers are also en- ged in production of necessary rectifier- e locomotives. The experience of the ench National Railways with this sys- n of electrification,[1] coupled with theoretical studies which have been de in the United States,[2,3] serve to ove doubts as to this system's prac- lity.

Test results, covering unbalanced load- of turbine generators, suggest that system could have application gen- erally in the United States, and partic- ularly in regions where there exists a power-generating system of high capacity, interconnected by a high-voltage trans- mission grid, which can absorb a minimum phase unbalance without difficulty.[2] The northeastern sector of the United States, together with areas in the southwest and northwest, has such a system of power generator and transmission. This sector covers the territory roughly north of the southern limit of the Tennessee Valley Authority area and the territory east of the Mississippi River. It practically coincides with the ICC Eastern District of Class I railways.

Communication

Automation can be used effectively to reduce operating cost. As applied to railways, in addition to the modern push- button classification yards, this involves centralized data processing and elec- tronic computer control of such things as payrolls, waybills, invoices, inventory, reservations, accounting data, train make- up, and movements of trains and cars. Transmission of large quantities of data in digital form is required from remote points to the central computer center, with reversed transmission of instruc- tions, again in digital form, from the center to remote points. Envisioned for the future are systems of automatic train operation under control of a central computer center, or central brain, re- quiring continuous flow of feedback in- formation from trains to the central brain and a reverse flow of instructions from the central brain to trains.

Moreover, the concept of "real-time" operation of railroads presupposes the

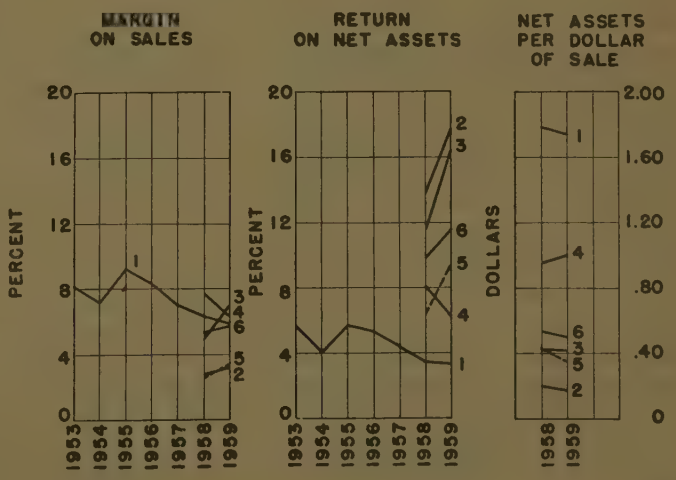


Fig. 3. Economic results of transportation company operations

1—Class I railroads
2—Motor carriers
3—Interstate bus companies
4—Inland water carriers
5—Air transport
6—Total manufacturing (for comparison)

ability to determine in advance the various movements of trains and cars required to accommodate the ebb and flow of traffic to be delivered from or to a given railroad. The production of this predetermined information requires complete use of data processing and computer facilities.

Practical automation of railway operating functions calls for a communication system that will provide many digital transmission channels and a limited number of voice channels, at low cost per channel. A microwave system for through-communication would meet this requirement. One that would be suitable for many railroads provides 30 voice channels, each of which can be subdivided into 15 digital channels. This gives a possible maximum of 450 digital channels, or say 10 voice channels and 300 digital channels. Initial capital investment and maintenance and operating costs are low as compared with any other system. Microwave is not affected by inductive interference from the traction electric power supply system. Relay stations could be located about 30 miles apart, which roughly coincides with the desired spacing of power supply points for the electrification system. For through-transmission of information, microwave appears to be ideal as a part of the technological package.

There remains, however, the problem of a local communication system which will make possible end-to-end communication on trains, between trains, between trains and wayside stations, and through communication between a control center and wayside stations, and between the control center and trains. Very-high-frequency (vhf) radiobroadcasting from microwave relay stations could be used. A more desirable system, however, would be the use of vhf signals, carried by the conductors of the electrification power supply system, and inductively coupled to moving trains and wayside locations. There are two known systems of this type, both of which probably would require further development for the proposed application. Any such system must provide numerous digital channels and a few voice channels. It should also make possible the bridging of necessary emergency through communication channels if there is a loss of one microwave repeater station.

A high-capacity communication system, having a low initial capital cost and free of inductive interference from the electrification power system, can be provided, and it should produce maintenance and operating savings over existing systems sufficient to justify its initial cost. A microwave system, coupled with a vhf local system, appears proper for inclusion in the technological package.

Centralized Traffic Control

Taken as a group, railroads in the northeast sector of the United States have more track miles than needed to handle their maximum traffic. Consequently they are burdened with excessive cost in the maintenance of way accounts. Centralized traffic control (CTC) is a well-known system by which train movement can be controlled from a single location. A dispatcher at that point can operate track switches and wayside signals up to distances as remote as 600 miles from his office. This system will allow elimination of trackage and of manually controlled interlockers, thereby producing savings on maintenance and operating costs.

In general, a single-track railroad with CTC has 80% the capacity of a 2-track railroad without it. In many cases, a 2-track road can be reduced to a single-track by installing CTC, and it will still have adequate capacity for handling existing traffic. As will be seen later, electrification will increase the speed of train movement, which will permit reduction in trackage without decreasing capacity. Similarly, 4-track and 3-track railroads can have less track mileage. Reduction in the number of tracks also makes possible the establishment of a work roadway for off-track maintenance. The resulting savings will justify the capital expenditure for the installation of CTC. Furthermore, the reduction also cuts the cost of electrification. CTC requires new signaling of the railroad, but its cost is justified by the savings produced. In this way, the signal changes which ordinarily would result from electrification are not burdensome.

CTC qualifies for inclusion in the package because it reduces trackage and produces permanent maintenance and operating savings. When it is combined with higher train speeds resulting from electrification, equivalent or increased track capacity is provided. Off-track maintenance, made possible by eliminating excessive tracks, is another advantage.

Converted Electric Locomotives

It is generally considered, at the average annual usage made of road diesel-electric locomotives by the Class I railroads of the United States, that their economic life is 12 to 15 years. At the end of this period the prime mover and its auxiliaries must be replaced. However, the mechanical parts of the loc motive, the traction motors, and the a brake system could be used economical for at least another 15 years as part of diesel locomotive. Basically, the hi rate of increase in maintenance cost as t locomotive ages is responsible for i short economic life. An electric loc motive has a rate of increase one-four or less than that of a diesel-electric loc motive.[3,4] Therefore, conversion to rectifier-type electric locomotive is pra tical, using the still-serviceable tracti motors, the mechanical parts, and t air-brake system of the diesel.

By converting these locomotives, in t general range of heavy freight train oper tion, gross ton-miles per train ho can be nearly doubled, provided spe is not restricted by track curvature other interference. This is because t electric locomotive is not limited by constant horsepower, as is the diesel an within the capacity of the tracti motors, can exert its maximum and co tinuous tractive efforts up to much high speeds. The tractive effort at the ma mum locomotive speeds will not dif from those of the locomotive as a dies electric unit. Due to the absence engine vibration and oil contaminat the mechanical parts will last long and conversion will cost less than a n diesel-electric unit.

Conversion of diesels to straight elec locomotives qualifies as part of the tec nological package because it will res in enormous maintenance savings, pr vide longer engine life, and, in most cas will add marginal savings through creased train speeds and possibly lo energy cost. It will conserve the use parts of old diesels and thereby the cost of locomotive replacement t minimum. Also, more attractive ser will ensue.

Electrification System

The system to be included in technological package is the single-ph commercial-frequency system, using 000 volts on the distribution system the tracks. Single-phase power is ta directly from the 3-phase high-vol network with alternate phases suppl different sections of the railway in or roughly to balance the phase load Studies made in this country and the perience gained from electrification France indicate that the residual p unbalance probably will not exceed limit of tolerance for larger power tems.[1,2] However, individual studies be necessary in each case. If requ known remedial measures can be app

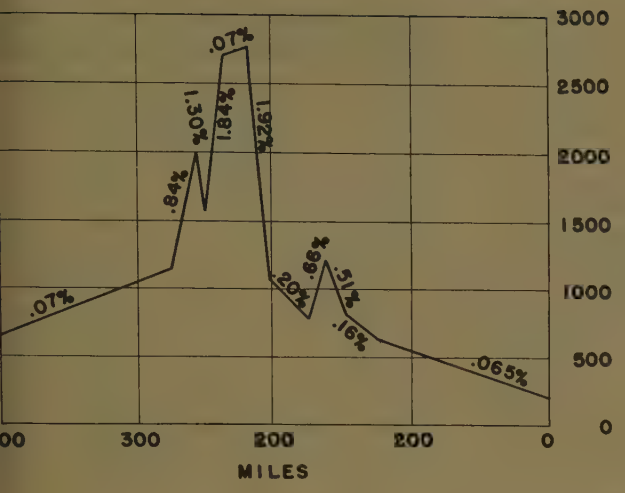


Fig. 4. Assumed average profile

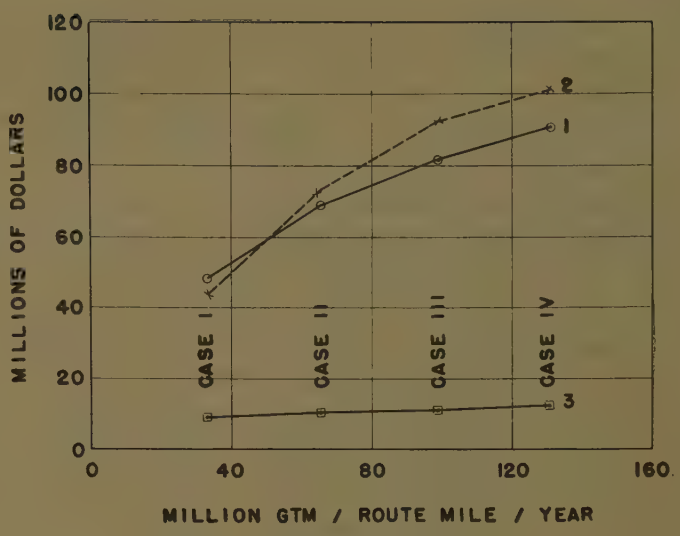


Fig. 5. Estimated net investment

1—Package with electrification
2—Package with dieselization
3—Communication and CTC, signal portion of package

As has been indicated, the central sta-
n industry may be willing to sell 25,000-
lt power to the railway at the right of
y, thus eliminating investment in
nsmission lines and stepdown sub-
tions as part of the electrification sys-
n. Comparatively good load factors
d diversity factors with existing de-
nds make the railway load attractive
r addition to central-station loading.
xty-cycle commercial frequency makes
ssible transformers of optimum size and
st as applied in the power stepdown
bstations and on the rectifier-type loco-
otives.

The 25,000-volt distribution voltage
duces the equivalent copper and the
ight of the catenary system for normal
ltage regulation and increases spacing
the power substations. The reduced
tenary weight makes possible smaller
d lighter poles and foundations, and a
nsequent over-all reduction in system
sts. With power feeds at 25- to 30-mile
acing, it is proposed that master
eakers at stepdown stations would con-
ol the entire section between feedpoints
order to remove power in the event of a
ort-circuit fault. Disconnect switches
intermediate sectionalizing stations
uld localize the fault; then power
uld be reapplied by closing the master
eakers. This could be done in a few
conds without seriously disturbing train
eration. For heavily loaded or un-
ly long sections, auxiliary feeders could
used, or a 3-wire system employing
ermediate autotransformers with track
l as the mid-point between feeders and
lley.

This electrification system qualifies for
lusion in the technological package
that its first cost will be less than
y other known system, and its operating
d maintenance cost will be relatively
v.

Economic Study

The economic advantages of the proposed technological package have been determined through a comparative economic study of all major costs, both with dieselization and with electrification. The existing cost level was determined for the year preceding conversion, and annual costs then derived for each of the 12 years following. The immediate effects of conversion by installation of the technological package could thus be shown.

Bases

The generalized study was based upon data, at a constant 1958 dollar value, from ICC published information and from other sources, applying to a typical Eastern District railroad operating from the eastern seaboard inload. Data were developed for four traffic levels utilizing a given railroad plant layout. The general bases of the study are as follows:

1. The existing railroad plant layout consists of a 2-track railroad having an average profile as shown in Fig. 4. There is some third track on the heavy-grade sections to accommodate movements of pusher locomotives. It was assumed that by use of CTC this basic plant could be converted to a single-track railroad with relatively frequently spaced passing sidings. The average spacing of these sidings varied with traffic level.

2. Operating data were developed on the basis of standard freight trains having the same trailing tonnage and the same number of road locomotives. All Eastern District railroads operating inland from the eastern seaboard have a considerably higher volume of revenue tonnage eastbound than westbound. This characteristic was reflected in the operating data in that fewer trains with lower trailing tonnage, but with the number of locomotive units required to balance the movement of motive power, made up the westbound train movement. Gross tons per train and the number of road and pusher locomotives were determined on the basis of ruling grades instead of the average grades which are shown in Fig. 4.

3. For simplification, it was assumed that a large fleet of diesel-electric locomotives, of 1,250 rhp (rail horsepower) rating, all 19 years of age, was available for electrification. The existing locomotives would be operating in their last depreciable year and at maximum maintenance cost. This cost was not based on system fleet averages; the average cost at the 16th year of life was used, as shown by published studies.[3] It was also assumed that the mechanical parts, traction motors, and air-brake equipment of the converted locomotive would have a life of at least 15 years, at which time the packaged electric equipment could be used to convert another fully depreciated diesel-electric locomotive.

4. Another assumption was that 30% of the capital cost for installation of a new communication system and the required CTC system would be met from depreciation funds on hand and the remaining cost from funds derived from the sale of bonds on the open market. It was further assumed that existing diesel-

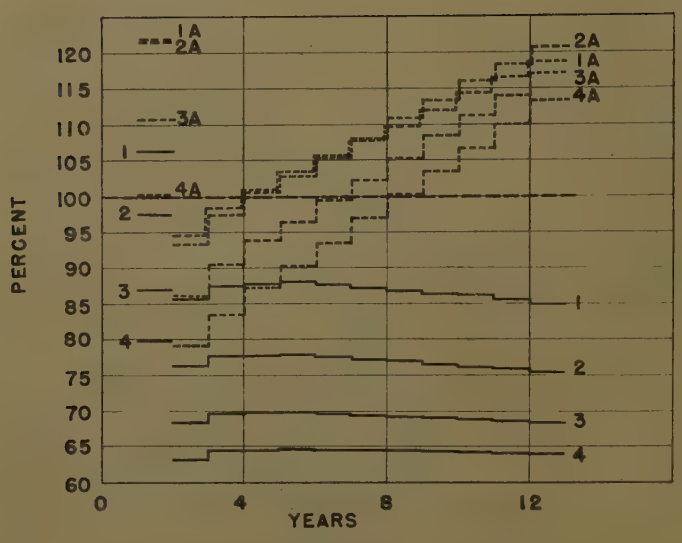


Fig. 6. Comparison of estimated existing cost level (100%) with estimated annual costs, including payments on capital, after conversion

1, 2, 3, 4—Cases I, II, III, and IV, respectively, with electrification
1A, 2A, 3A, 4A—Cases I, II, III, and IV, respectively, with dieselization

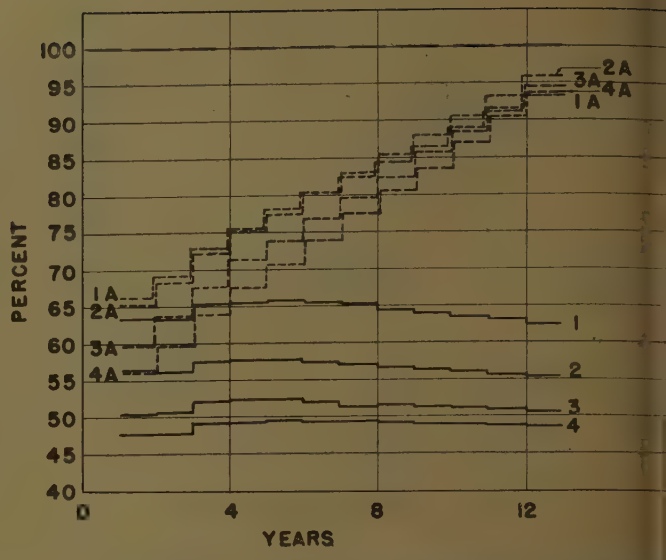


Fig. 7. Comparison of estimated existing cost level (100%) w estimated annual costs, not including payments on capital, after co version

1, 2, 3, 4—Cases I, II, III, and IV, respectively, with electrificat
1A, 2A, 3A, 4A—Cases I, II, III, and IV, respectively, with dies dieselization

electric locomotives would either be replaced by diesel-electric locomotives of a somewhat higher horsepower rating or would be converted to electric locomotives, and that the capital cost for this would be met by the sale of equipment trust certificates without recourse to depreciation funds. The cost of the new substations and distribution circuits required for electrification would be met by sale of bonds. In the case of both bond issues, sinking funds would be set up and the interest earnings of those funds would be credited against the interest cost on the bonds.

5. The existing cost level was established by determining charges for:

(*a*). Depreciation and property taxes on the existing communication and signal systems, together with depreciation for the last depreciable year on existing locomotives.

(*b*). Operation and maintenance of the existing communication and signal systems, maintenance of existing trackage, locomotive maintenance, engine house expense, and lubrication of locomotives.

(It will be noted that the existing cost level includes neither interest charges nor payments on capital.)

6. The total successive annual costs were established as the summation of:

(*a*). Sinking fund payments on bond issues and amortization payments on equipment trust certificates.

(*b*). Net interest on bond issues and interest on the outstanding balance for equipment trust certificates.

(*c*). Depreciation charges on the new communication and CTC systems, the electric power system, and new diesel-electric locomotives or the converted locomotives.

(*d*). Property taxes on the new communication, CTC, and electric power supply systems.

(*e*). Operating costs for the same items listed in (d).

(*f*). Maintenance costs for the same items as in (*d*) and (*e*) and for the revised track layout, and locomotives—either diesel-electric or converted electric—together with engine house expense and locomotive lubrication.

Unit values used in establishing both existing cost levels and annual co after conversion are believed to be co servative. They are available from author upon application.

ANALYSIS

Fig. 5 shows the required investm for the four cases studied, includ the proportion going for the new comm nication and CTC systems. These syste require a relatively constant investme whereas the totals show increases whi occur at decreasing rates for the heav traffic cases. In all except case I, total investment for the technolog package utilizing railway electrificatio

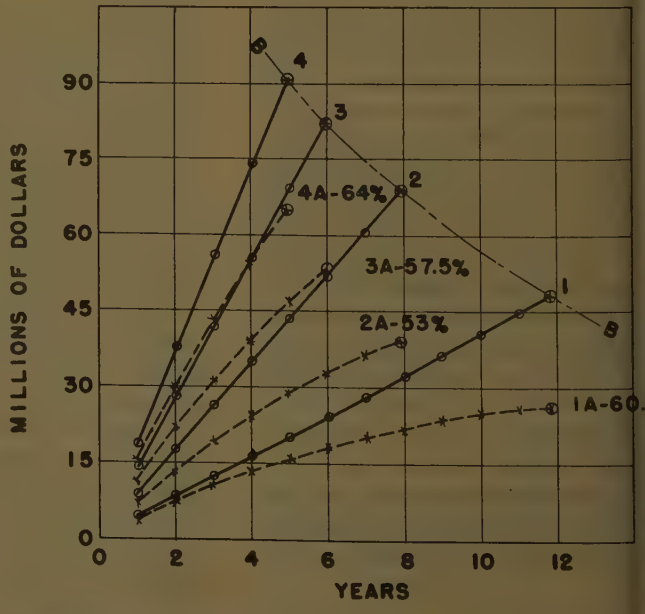


Fig. 8. Cumulative savings

1, 2, 3, 4—Cases I, II, III, and IV, respectively, with electrification. 1A, 2A, 3A, 4A—Cases I, II, III, and IV, respectively, with dieselization. B–B—Curve of 100% accumulation of net capital investment for package with electrification; percentage figures indicate portion of net capital investment which would accumulate in time shown for package with dieselization

Table I. Proportion of Package Represented by New Communication and CTC Systems

Cases	Type of Locomotive	Proportion of Capital Investment	Proportion* of 100% Accumulation
I	Diesel	20.0	0
	Electric	18.3	24.2
II	Diesel	14.2	0
	Electric	15.2	15.2
III	Diesel	12.1	0
	Electric	13.5	13.9
IV	Diesel	12.0	0
	Electric	13.7	12.1

* The dollar proportion in each case is approximately $11,000,000.

less than that for the package utilizing dieselization. This is attributed to lower cost of conversion as compared with purchasing new diesel-electric locomotives, and to a somewhat smaller fleet of electric locomotives. As the result of higher train speeds and consequent better locomotive utilization, fewer electric locomotives are needed.

One method of analyzing the results obtained by conversion is to compare annual costs, including payments on capital. Since dollar values are different for each case studied, Fig. 6 has been prepared to show the relationship of the four cases in percentages of the existing cost level. This chart clearly demonstrates that, except for the first year when large first payments were made, the technological package with electrification has permanently stabilized annual costs below the existing level.

The package with dieselization is markedly different in that, except for the first year, costs initially drop below but not as far below the existing level as in the case of electrification, and then rapidly rise to well above that level. Costs below the existing level mean increased profits, while those above it signify decreased profits. The striking difference in the two packages arises from comparatively rapid rise in diesel maintenance costs, coupled with greater charges for financing and depreciation which result from higher unit cost per locomotive, shorter economic life, and a larger required fleet.

The economic results of conversion by the technological package can be shown by studying the existing cost level relationship to successive annual costs after conversion, but not including payments of capital; see Fig. 7. The difference between the two costs represents savings which could be effected and which can be expressed cumulatively in terms of dollars for each case. This was done in Fig. 8.

This chart shows the basic divergence in trends in a different manner. In each case, the rate of rise in cumulative savings is greater, and practically uniform, for the package with electrification, but has a drooping characteristic at a lower level for the package with dieselization. In each of the four cases, 100% accumulation of total capital costs for the technological package with electrification is achieved long before such accumulation for the package with dieselization would be attained.

Division of the technological package between new communication and CTC systems and motive-power requirements is indicated in Table I. The communication and CTC portion of the package is relatively small in both capital investment and cumulative savings. The bulk of capital investment and cumulative savings are accounted for by the motive-power portion of the package in each case studied. In this connection, the capital cost of the electrification substation and distribution systems accounts for 40% in case I and 23% in case IV of the total required motive-power investment.

The new diesel-electric locomotives were assumed to have a moderately higher horsepower rating than those previously used. By converting existing diesels to electric locomotives, the continuous rhp would be more than doubled because, at the low speeds at which the continuous tractive effort of the diesel-electric locomotive is exerted, the traction motors are grossly underloaded as to applied voltage. Conversion to electric locomotives would make possible full-capacity use of traction motors at the continuous rating. These increased continuous horsepower ratings mean higher average train speeds, as indicated in Table II.

Electrification will make possible an average round-trip speed, neglecting stops and turn-around time, 25% greater than obtained with existing diesels. The full advantage of the electric locomotive is brought out in the data showing the average speed ascending the 1.92% grade, where use of electric locomotives would increase the speed up the grade approximately 2.5 times more than that now obtained.

Discussion of Results

The economic advantage of installing the technological package with electrification of the railroad is clearly indicated for the conditions assumed. Even when payments on capital are included, there is a reduction of cost below the prevailing level and in all cases the level remained approximately constant over the 12-year period. In the case of the technological package with dieselization, there is also a drop below the existing cost level, but to a lesser degree, followed by a pronounced upward trend in the annual cost. Therefore, when payments on capital are included, the annual cost level soon exceeds the existing one. This is true for all four cases studied.

The study also indicates that cumulative savings for the package with electrification would equal 100% of the net capital investment in 5 to 12 years, depending upon the traffic level. Cumulative savings for such periods materially exceed those of the package with dieselization.

The results of a general study of this nature can be altered by refinement of the estimates as to (1) capital investment cost required, (2) existing cost level to be used as a reference, (3) annual costs which will be incurred after installation of the technological package, and (4) operation over different profiles. Other cases could be studied involving an existing single-track, 3-track, or 4-track railroad, with average profiles and traffic levels differing from those used in this case. Also, definitive studies can be made of individual situations, using factual data rather than estimates.

This analysis shows, however, that in all studies involving relatively heavy traffic the same trends will prevail when comparing dieselization and electrification. The effects of lower capital costs, smaller locomotive fleet, longer economic life, and radically lower rate of rise in maintenance costs for electric locomotives overwhelmingly dictate a lower and stabilized cost level as compared with that of the technological package with dieselization. In all such cases, an economic advantage will be realized by installing the package with railway electrification. The rate of rise of cumulative savings and the time periods required for 100% accumulation of capital requirements will vary, but the basic relationship between the packages will be the same.

Table II. Train Speeds and Locomotive Horsepower

Type of Locomotive	Average Round-Trip Speed, %	Speed Ascending 1.92% Grade, %	RHP per Locomotive, %
Existing diesel-electric	100.0	100.0	100.0
New diesel-electric	107.5	122.0	116.0
Converted electric	125.0	256.0*	212.0

* At maximum short-time horsepower rating.

Conclusions

This study demonstrates that the installation of the technological package with railway electrification would meet the basic needs of the railroads in the Eastern District, as stated previously, by reduction in operating and maintenance costs with a minimum of capital expenditure. Also, to assist in attracting additional traffic, greatly increased speed for freight trains would be required.

References

1. French Technical Advances in the Field of Railroad Electrification, F. Nouvion. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 79, Sept. 1960, pp. 241–48.

2. Technical Aspects of Providing Service to Single-Phase 60-Cycle Railroad Loads, T. J. Nagel, A. F. Gabrielle. *Ibid.*, vol. 77, July 1958, pp. 172–76.

3. Locomotive Repair Costs and Their Economic Meaning to the Railways of the United States, H. F. Brown. *Ibid.*, pt. I (*Communication and Electronics*), vol. 80, Sept. 1961, pp. 209–16.

4. Virginian Railway Motor-Generator Electric Locomotive Maintenance Costs, T. F. Perkinson. *Ibid.*, pt. II (*Applications and Industry*), vol. 79, Mar. 1960, pp. 33–35.

Discussion

L. W. Birch (Ohio Brass Company, Mansfield, Ohio): In my old psychology book there are two types of escape mechanisms, vividly described, that stand out in my memory. The first escape mechanism deals with the "scapegoat" and the second with projection—accusing the other fellow of your own faults. Why have I mentioned these two escape mechanisms? I have done this because the overhead catenary distribution systems are usually considered extremely costly and have, many times, been offered as a reason why an attempt to electrify a railroad has failed. Let us see if this reasoning is right.

Some readers may be familiar with the work of the Joint Committee on Railroad Electrification. This committee, which consisted of members from railroads, the coal industry, power companies, cable manufacturers, electric equipment manufacturers, and railway consultants, was set up to study the future of railroad electrification in the United States and Canada. The committee, active in the early 1950's, spent almost 2 years in this study. Three contact voltages were involved: (1) 3,000 volts direct current; (2) 12,000 volts alternating current; and (3) 25,000 volts alternating current, mentioned in the paper.

In 1950 and 1951, the use of 25-kv trolley voltage was foreign to most of those participating in the study. However, a few of us had helped design and install the 17-mile electrification of Henry Ford's Detroit, Toledo & Ironton Railroad in 1925–26. The trolley voltage of this system was 22 kv. In 1944–45, several of the Land Transportation Committee members worked on a New York Central estimate, involving electrification of the Central from Harmon to Buffalo, N. Y. This study was based on the same trolley voltages that had been considered in 1950–51 by the Joint Committee on Railroad Electrification.

Let us consider the New York Central costs and the Joint Committee's costs for 25-kv overhead, not including certain standard percentages such as contractor's profits, insurance, and interest. We will compare only the catenary overhead system. If the 1950 cost index for the 25-kv catenary overhead, without transmission lines, is 100, then the 1951 Joint Committee's index for the same voltage and the same system is 84, and a review of today's costs indicate that the cost index for a 25-kv catenary system is approximately 130—not an excessive increase over the 1945 index.

It is doubtful if the overhead cost increases are as high, percentagewise, as some of the other costs such as equipment or signaling. Also, there have been many improvements in the distribution art since 1945. Based on these indexes for catenary overhead distribution costs, there is little cause to ask, "Are the overhead distribution costs retarding railway electrification?"—the subject of a paper which I presented at the 1948 AIEE Winter General Meeting.

B. A. Ross (American Electric Power Service Corporation, New York, N. Y.): The author's analysis of the economics of a proposed railroad improvement package, including commercial-frequency electrification, is both interesting and timely.

The utility industry is well able to handle the railroad loads with a minimum of new facilities. Studies indicate that existing facilities can, in most areas, tolerate the phase unbalances which might result from the railway loads. While the utility can deliver energy at the catenary voltage as indicated, it should be remembered that, in most areas, only the major subtransmission circuits (i.e., those circuits in the 100- to 150-kv voltage range) could tolerate the phase unbalances of the single-phase railway load without affecting the equipment of other utility customers. This would probably require, in some areas, limited extensions in the utility transmission system, or slight changes in the number and location of the railway substation facilities.

An additional economic factor favoring electrification is the ability of the utility industry to hold, or perhaps even decrease, the cost of electric energy, whereas future oil prices will, in all probability, continue to rise. The electric utility's ability to control its future energy prices is influenced by the fact that it can: (1) utilize the most economic energy source or combination of sources; (2) employ either conventional or unconventional methods of energy conversion with little restriction as to plant location or size; (3) still further increase the efficiency of energy conversion by using larger generating units, new methods of conversion, and improved techniques; and (4) improve over-all electric system efficiency and reliability by utilizing higher transmission voltages, higher capacity, and more efficient station equipment, as well as improved relaying and operating methods.

H. C. Cross: In making the economic study referred to by Mr. Ross, the element of power cost was omitted, since today's cost of electrical and diesel fuel, delivered at the locomotive, are considered roughly equal. Significant differences may be found in localized geographical areas. Since the electric utility industry is able to control future energy prices, this element should have thorough consideration when a definitive feasability study is made. Also noteworthy is Mr. Ross' opinion that phase unbalance can be tolerated in most areas, thus assuming the industry's capability of handling "the railroad loads with a minimum of new facilities."

The remarks by Mr. Birch indicate that the estimated cost of a 25-kv catenary system has not increased during the last 15 years in proportion to other elements of electrification cost. In this study, the estimate is approximately the same as that of the CTC system on a track-mile basis. For a project such as assumed, the reduction in track miles by use of CTC eliminates approximately 35% of the catenary system cost required for a 2-track railroad. Perfection of off-track installation methods might provide reductions in unit costs below the estimates.

In addition to cost, the fixed property investment required for a catenary system has deterred adoption of electrification. However, today's railroads are making just such investments by adopting CTC in order to obtain operating and maintenance savings. The over-all savings resulting from use of the technological package adequately justify the catenary system.

The most significant factor involved in the study is maintenance cost of diesel-electric locomotives. Following presentation of the paper, the author has received an indication that railway management may not be aware of the extent to which this cost rises with advancing age of their locomotives. Close study of this trend by individual railroads probably will disclose, in many cases, a steeper rate of rise than that estimated in the study. This factor is crucial with respct to economic justification of railway electrification.

On the Optimum Synthesis of Random Sampling Multipole Filters with Stationary Inputs

H. C. HSIEH
NONMEMBER AIEE

IN MANY practical engineering problems the sampling intervals for the system are not fixed. They are actually random variables. For example, in radar with nominal scanning period T, occasionally some samples of the return signals are absent because of noise or other interference. It is referred to as the "missed sample" problem in radar.

This kind of system has been studied to a certain extent by several previous investigators. Kalman considered the optimum synthesis of a random sampling system with step input and quadratic error criterion.[1] His synthesis procedure is based on the technique of dynamic programming. He made detailed studies on the stability of such systems.

Bergen investigated the minimum mean-square synthesis of a randomly sampled system with one stationary stochastic input.[2] The evaluation of power spectral density after sampling in terms of that before sampling and the generating function of the random sampling process has been obtained. His synthesis procedure is based upon the standard Wiener's spectral-density factorization technique.[3] It has been pointed out that, with the assumption of rational functions in s for the spectral densities before sampling, straightforward factorization of the spectral densities after sampling can be carried out for two common and important classes of sampling.

The purpose of this paper is to obtain the design equations for an optimum random sampling multipole filter with n stationary inputs and m outputs. This optimum system is also optimum in the sense of Wiener. The system is assumed to be linear and time invariant. A set of integral equations which the optimum weighting functions must satisfy is obtained. Transform method is then used to solve these equations.

Formulation of the Filtering Problem

Description of Random Sampling Processes and System Inputs

The random sampling of an input series means that the samples of this input are taken at random time t_i. The ith interval T_i between the sampling times t_i and t_{i+1} is defined as:

$$T_i = t_{i+1} - t_i \tag{1}$$

Instead of working directly with the sampling time t_i, it is convenient to assume that these sampling intervals T_i are statistically independent random variables with a first-order probability density function $f_1(T)$. In other words,

$$Pr(T < T_i < T + dT) = f_1(T)dT \text{ for all } i \tag{2}$$

This density function is either continuous or discrete.

Let the time series $u(t)$ be a sequence of very narrow pulses with width γ and of unity amplitude, which occur at random times. The intervals T_k between the leading edges of successive pulses are statistically independent variables with the same first-order probability density function $f_1(T)$. This sampling process is thus stationary. The random pulse modulation of the input may then be described by

$$i_u(t) = \frac{1}{\bar{u}} u(t) i(t) \tag{3A}$$

Here $\bar{u}$ is the ensemble average of $u(t)$ and is arbitrarily introduced in the equation in order to simplify the subsequent manipulations. This sequence of random pulse modulation of the input is shown in Fig. 1. The randomly sampled input $i^*(t)$, which is usually regarded as the impulse modulation of the continuous input can then be obtained as the pulse width approaches zero. Hence:

$$i^*(t) = \lim_{\gamma \to 0} \frac{1}{\bar{u}} u(t) i(t)$$

$$= \frac{1}{\bar{u}} u^*(t) i(t) \tag{3B}$$

where the asterisk denotes the sampled quantity and the bar, the statistical average.

The system diagram is shown in Fig. 2. The input before sampling at each terminal of the multipole filter consists of stationary random signal and noise. Thus:

$$i_k(t) = S_k(t) + N_k(t) \tag{4}$$

These processes have zero means and known correlation functions. In dealing with multipole systems, it is reasonable to assume that the random sampling process of each terminal is statistically independent of those of the others and also of the input processes. Hence, for the pulse modulation:

$$i_{u_k}(t) = \frac{1}{\bar{u}_k} u_k(t) i_k(t) \tag{5}$$

and

$$\overline{u_{k'}(t) u_k(t+\tau)} = \overline{u_{k'}}\ \overline{u_k} \text{ for } k' \neq k$$

$$= \phi_{u_k u_k}(\tau) \text{ for } k' = k \tag{6}$$

where $\phi_{u_k u_k}(\tau)$ is the autocorrelation function of the sampling process. It can easily be shown that, with further assump-

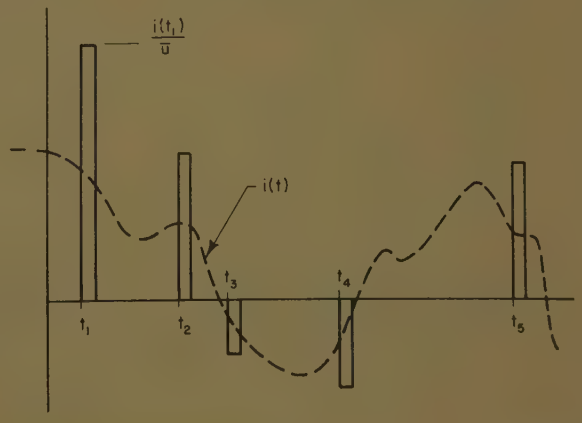


Fig. 1. Pulse sequence $i_u(t)$ of randomly sampled input

Paper 61-802, recommended by the AIEE Basic Sciences Committee of the Science and Electronics Division and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Ithaca, N. Y., June 18–23, 1961. Manuscript submitted February 6, 1961; made available for printing April 7, 1961.

H. C. Hsieh is with the University of California, Los Angeles, Calif.

The author expresses his gratitude to Dr. C. T. Leondes of the University of California at Los Angeles, for his many helpful suggestions. This research was supported by the U. S. Air Force under Contract no. *AF49(638)-438*, monitored by the Air Force Office of Scientific Research of the Air Research and Development Command.

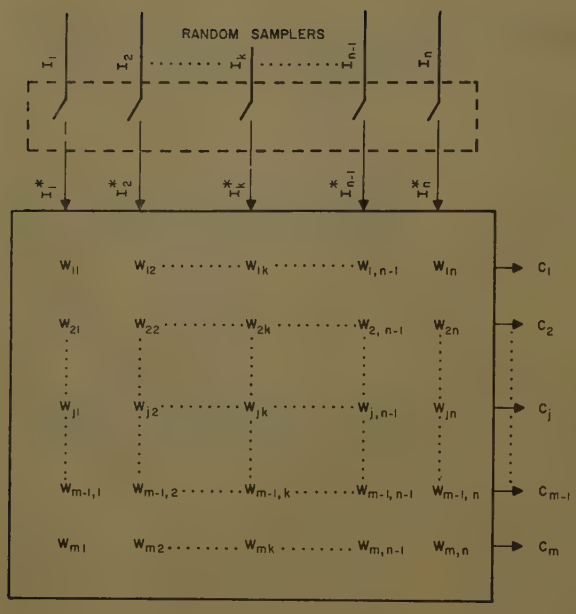


Fig. 2 (left). Block diagram for n+m pole filter

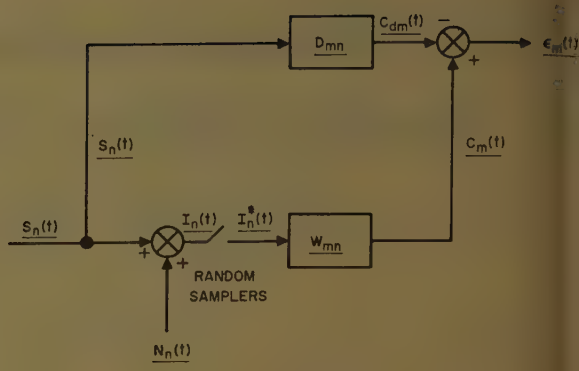


Fig. 3 (right). Error generation diagram

tion of ergodicity for the process, the average value of $u_k(t)$ is:

$$\bar{u}_k=\frac{\gamma}{\bar{T}_k} \tag{7}$$

where $\bar{T}_k$ is the average sampling period. The derivation of $\phi_{u_k u_k}(\tau)$ and its limiting case $\phi_{u_k^* u_k^*}(\tau)$ is given in reference 2. The latter is expressed by:

$$\phi_{u_k^* u_k^*}(\tau)=\bar{u}_k\gamma\Big[\delta(\tau)+\sum_{n=1}^{\infty}\underbrace{\int\cdots\int}_{n\text{ fold}}\times \delta(\tau-T_1-\ldots-T_n)f_{nk}(T_1\ldots T_n)\,dT_1\ldots dT_n\Big] \quad \text{for } \tau\geqq 0 \tag{8}$$

where f_{nk} is the nth order probability density function.

Derivation of System Equations

In deriving the system equations, pulse modulation is used at the beginning. The system errors are defined, as usual, as the difference between actual outputs and desired outputs; see Fig. 3. Thus:

$$\begin{aligned}\epsilon_j(t)&=C_j(t)-C_{dj}(t)\\&=\sum_{k=1}^{n}\int_0^{\infty}W_{jk}(\tau)i_{u_k}(t-\tau)d\tau-\sum_{k=1}^{n}\int_{-\infty}^{\infty}D_{jk}(\tau)S_k(t-\tau)d\tau\\&=\sum_{k=1}^{n}\int_0^{\infty}W_{jk}(\tau)\Big[S_k(t-\tau)+N_k(t-\tau)\Big]\frac{u_k(t-\tau)}{\bar{u}_k}\,d\tau-\sum_{k=1}^{n}\int_{-\infty}^{\infty}D_{jk}(\tau)S_k(t-\tau)d\tau\end{aligned} \qquad j=1,2,\ldots,m \tag{9}$$

and

$$\begin{aligned}\overline{\epsilon_j^2}=&\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_0^{\infty}W_{jk'}(\tau_2)d\tau_2\times\int_0^{\infty}W_{jk}(\tau_1)\phi_{i_{k'}i_k}{}^{u^2}(\tau_2-\tau_1)d\tau_1-\\&2\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_0^{\infty}W_{jk'}(\tau_2)d\tau_2\times\int_{-\infty}^{\infty}D_{jk}(\tau_1)\phi_{i_{k'}s_k}{}^{u}(\tau_2-\tau_1)d\tau_1+\\&\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_{-\infty}^{\infty}D_{jk'}(\tau_2)d\tau_2\times\int_{-\infty}^{\infty}D_{jk}(\tau_1)\phi_{s_{k'}s_k}(\tau_2-\tau_1)d\tau_1\end{aligned} \qquad j=1,2,\ldots,m \tag{10}$$

where

$$\phi_{i_{k'}i_k}{}^{u^2}(\tau)=\phi_{s_{k'}s_k}{}^{u^2}(\tau)+\phi_{s_{k'}N_k}{}^{u^2}(\tau)+\phi_{N_{k'}s_k}{}^{u^2}(\tau)+\phi_{N_{k'}N_k}{}^{u^2}(\tau)$$

$$\phi_{i_{k'}s_k}{}^{u}(\tau)=\phi_{s_{k'}s_k}{}^{u}(\tau)+\phi_{N_{k'}s_k}{}^{u}(\tau)$$

are the various correlation functions. The synthesis criterion is, then, to minimize these system mean-square errors.

Investigation is now made of all the correlation functions after sampling. It has been assumed that each of the sampling processes is statistically independent of the others and also of the input processes.

Thus:

$$\begin{aligned}\phi_{s_{k'}s_k}{}^{u^2}(\tau)&=\overline{S_{k'}(t)\frac{u_{k'}(t)}{\bar{u}_{k'}}S_k(t+\tau)\frac{u_k(t+\tau)}{\bar{u}_k}}\\&=\frac{1}{\bar{u}_k{}^2}\phi_{u_k u_k}(\tau)\phi_{s_k s_k}(\tau)\ \text{for } k=k'\\&=\phi_{s_{k'}s_k}(\tau)\ \text{for } k\neq k'\end{aligned} \tag{11}$$

Similarly

$$\phi_{s_{k'}N_k}{}^{u^2}(\tau)=\frac{1}{\bar{u}_k{}^2}\phi_{u_k u_k}(\tau)\phi_{s_k N_k}(\tau)\ \text{for } k=k'$$

$$=\phi_{s_{k'}N_k}(\tau)\ \text{for } k\neq k' \tag{12}$$

On the other hand,

$$\phi_{s_{k'}s_k}{}^{u}=\frac{\overline{S_{k'}(t)u_{k'}(t)S_k(t+\tau)}}{\bar{u}_{k'}}$$

$$=\phi_{s_{k'}s_k}\quad \text{for all } k \text{ and } k' \tag{13}$$

and

$$\phi_{N_{k'}s_k}{}^{u}(\tau)=\phi_{N_{k'}s_k}(\tau)\quad \text{for all } k \text{ and } k' \tag{14}$$

Hence the following equations are obtained:

$$\phi_{i_{k'}s_k}{}^{u}(\tau)=\phi_{i_{k'}s_k}(\tau)\quad \text{for all } k \text{ and } k' \tag{15}$$

and

$$\phi_{i_{k'}i_k}{}^{u^2}(\tau)=\phi_{i_{k'}i_k}(\tau)\quad \text{for all } k\neq k' \tag{16}$$

$$=\frac{1}{\bar{u}_k{}^2}\phi_{u_k u_k}(\tau)\phi_{i_k i_k}(\tau)\ \text{for } k=k'$$

In other words, the cross-correlation functions of the sampled inputs and the cross correlation functions of the sampled and continuous inputs, are unaffected by the sampling processes. Since these correlation functions do not depend on the nature of the random-pulse trains, it is evident that, as the pulse width τ approaches zero, the following results are true:

$$\phi_{i_{k'}s_k}{}^{*}(\tau)=\lim_{\gamma\to 0}\phi_{i_{k'}s_k}{}^{u}(\tau)=\phi_{i_{k'}s_k}(\tau)$$

for all k and k' (17)

and

$$\phi_{i_{k'}i_k}{}^{*}(\tau)=\lim_{\gamma\to 0}\phi_{i_{k'}i_k}{}^{u^2}(\tau)=\phi_{i_{k'}i_k}(\tau)$$

for all $k'\neq k$ (18)

The autocorrelation functions of the sampled inputs are, on the other hand, affected by the sampling processes. Thus as the pulse width approaches zero we have:

$$\phi_{i_k i_k}{}^{*}(\tau)=\lim_{\gamma\to 0}\phi_{i_k i_k}{}^{u^2}(\tau)$$

$$=\frac{1}{\bar{u}_k{}^2}\phi_{u_k^* u_k^*}(\tau)\phi_{i_k i_k}(\tau) \tag{19}$$

The mean-square errors are, in the limiting case:

$$\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_0^\infty W_{jk'}(\tau_2)d\tau_2\times$$
$$\int_0^\infty W_{jk}(\tau_1)\phi_{i_k'i_k}{}^*(\tau_2-\tau_1)d\tau_1-$$
$$2\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_0^\infty W_{jk'}(\tau_2)d\tau_2\times$$
$$\int_{-\infty}^\infty D_{jk}(\tau_1)\phi_{i_k'S_k}(\tau_2-\tau_1)d\tau_1+$$
$$\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_{-\infty}^\infty D_{jk'}(\tau_2)d\tau_2\times$$
$$\int_{-\infty}^\infty D_{jk}(\tau_1)\phi_{S_k'S_k}(\tau_2-\tau_1)d\tau_1$$
$$j=1, 2, \ldots, m \quad (20)$$

he necessary and sufficient condition for btaining minimum errors is that the sys-m weighting functions must satisfy the llowing integral equations:

$$\sum_{k=1}^{n}\int_0^\infty W_{jk}(\tau_1)\phi_{i_k'i_k}{}^*(\tau_2-\tau_1)d\tau_1$$
$$=\sum_{k=1}^{n}\int_{-\infty}^\infty D_{jk}(\tau_1)\phi_{i_k'S_k}(\tau_2-\tau_1)d\tau_1$$
$$\text{for } \tau_2\geqq 0$$
$$k'=1, 2, \ldots, n$$
$$j=1, 2, \ldots, m \quad (21)$$

hese are the generalized Wiener-Hopf quations for a random sampling system. he procedure used in deriving these quations may be found elsewhere.[4,5] The minimum mean-square errors are:

$$\overline{\epsilon_j^2}_{\min}=\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_{-\infty}^\infty D_{jk'}(\tau_2)d\tau_2\times$$
$$\int_{-\infty}^\infty D_{jk}(\tau_1)\phi_{S_k'S_k}(\tau_2-\tau_1)d\tau_1-$$
$$\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_0^\infty W_{jk'}(\tau_2)d\tau_2\times$$
$$\int_0^\infty W_{jk}(\tau_1)\phi_{i_k'i_k}{}^*(\tau_2-\tau_1)d\tau_1$$
$$=\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_{-\infty}^\infty D_{jk'}(\tau_2)d\tau_2\times$$
$$\int_{-\infty}^\infty D_{jk}(\tau_1)\phi_{S_k'S_k}(\tau_2-\tau_1)d\tau_1-$$
$$\sum_{k'=1}^{n}\sum_{k=1}^{n}\int_0^\infty W_{jk'}(\tau_2)d\tau_2\times$$
$$\int_{-\infty}^\infty D_{jk}(\tau_1)\phi_{i_k'S_k}(\tau_2-\tau_1)d\tau_1$$
$$j=1, 2, \ldots, m \quad (22)$$

inimization Equations for Optimum Filters

OWER SPECTRAL DENSITIES AFTER RANDOM SAMPLING

The integral equations for the optimal ighting functions can be solved as usual by employing the transform method. Thus the transformation of the correlations should first be discussed. As was shown in the previous section, the cross-correlation functions are not affected by the sampling process. Therefore, the cross spectral densities, which are the transforms of the cross-correlation functions, are the same before and after the sampling.

The autocorrelation function of the input after sampling is proportional to the product of the autocorrelation function of that input before sampling and the autocorrelation function of the sampling process, as shown in equation 19. Since the autocorrelation function is an even function, it is possible to express it as:

$$\phi_{kk}{}^*(\tau)=\phi_{k_1k_1}{}^*(-\tau)+\phi_{k_1k_1}{}^*(\tau) \quad (23)$$

where

$$\phi_{k_1k_1}{}^*(\tau)=\phi_{kk}{}^*(\tau) \text{ for } \tau>0$$

and

$$\phi_{k_1k_1}{}^*(0)=1/2\phi_{kk}{}^*(0)$$

The two-sided Laplace transform of $\phi_{kk}{}^*(\tau)$ is then:

$$G_{kk}{}^*(s)=G_{k_1k_1}{}^*(s)+G_{k_1k_1}{}^*(-s) \quad (24)$$

where

$$G_{k_1k_1}{}^*(s)=L_1[\phi_{k_1k_1}{}^*(\tau)]$$

where L_1 denotes the conventional one-sided Laplace transform.

It has been shown by Bergen that $G_{k_1k_1}{}^*(s)$ can be expressed by a complex convolution integral as:

$$G_{k_1k_1}{}^*(s)=\frac{\bar{T}_k}{2\pi j}\int_{c-j\infty}^{c+j\infty}\frac{G_{k_1k_1}(\lambda)}{1-M_k(s-\lambda)}\,d\lambda \quad (25)$$

where $G_{k_1k_1}(s)$ is the one-sided Laplace transform of the autocorrelation function before sampling and

$$M_k(s)=\text{generating function of } \{T\} \text{ in complex variable } -S$$
$$=E[e^{-sT}]=L_1[f_{1k}(T)] \quad (26)$$

The function $G_{k_1k_1}(\lambda)$ will have singularities in the left half of the λ plane only and $1/1-M_k(s-\lambda)$ will have singularities in the right half of the λ plane only. There exists an analytic strip around the imaginary axis which separates the singularities of these two functions. The path of integration of the integral lies in this strip. The function $G_{k_1k_1}{}^*(s)$ is analytic in the right half of the s plane.

The spectral-density convolution integral of equation 25 can be evaluated by contour integration due to the analyticity of $G_{k_1k_1}(\lambda)$ and $1/1-M_k(s-\lambda)$ in the different halves of the λ plane.

Thus:

$$G_{k_1k_1}{}^*(s)=\frac{\bar{T}_k}{4\pi j}\int_{\Gamma_L}G_{k_1k_1}(\lambda)\frac{1+M_k(s-\lambda)}{1-M_k(s-\lambda)}\,d\lambda \quad (27)$$

where $[1+M_k(s-\lambda)/1-M_k(s-\lambda)]$ is analytic in the left half plane. The closed contour Γ_L includes the path parallel to the imaginary axis and a semi-infinite circle in the left half plane. Thus the only singularities enclosed by this contour are those of $G_{k_1k_1}(\lambda)$.

Several spectral densities after sampling which have been evaluated by Bergen, are given in the following paragraphs.[2]

Purely Random Sampling

If the sampling times are purely random, then the sampling intervals will be statistically independent with the exponential first-order probability density function.

$$f_{1k}(T)=\frac{1}{\bar{T}_k}e^{-T/\bar{T}_k} \qquad T\geqq 0$$
$$=0 \qquad T<0$$

Thus

$$M_k(s)=\frac{1}{s\bar{T}_k+1}$$

and

$$G_{k_1k_1}{}^*(s)=\frac{\bar{T}_k}{2\pi j}\int_{c-j\infty}^{c+j\infty}G_{k_1k_1}(\lambda)\times$$
$$\frac{1}{1-\dfrac{1}{(s-\lambda)\bar{T}_k+1}}\,d\lambda=G_{k_1k_1}(s)+\frac{\bar{T}_k\phi_{k_1k_1}(0)}{2}$$

Hence

$$G_{kk}{}^*(s)=G_{kk}(s)+\bar{T}_k\phi_{kk}(0) \quad (28)$$

The effect of purely random sampling is thus equivalent to adding white noise to the original spectral density.

Missed Samples Case

For a system with a nominal sampling interval of T_0, if it is assumed that the misses occur independently and that the probability of a miss is p_k, then:

$$f_{1k}(T)=\sum_{n=1}^{\infty}(1-p_k)p_k{}^{n-1}\delta(T-nT_0)$$

and

$$M_k(s)=\frac{(1-p_k)\,e^{-sT_0}}{1-p_ke^{-sT_0}}$$

It follows that:

$$G_{kk}{}^*(s)=\frac{T_0}{2\pi j}\int_{\Gamma_L}G_{kk}(\lambda)\times$$
$$\frac{1-e^{2\lambda T_0}}{[1-e^{-(s-\lambda)T_0}][1-e^{(s+\lambda)T_0}]}\,d\lambda+$$
$$\frac{p_kT_0}{1-p_k}\phi_{kk}(0) \quad (29)$$

For periodic sampling $p_k=0$, then:

$$G_{kk}(s)=\frac{T_0}{2\pi j}\int_{\Gamma L} G_{kk}(\lambda)\,\frac{1-e^{2\lambda T_0}}{[1-e^{-(s-\lambda)T_0}]\,[1-e^{(s+\lambda)T_0}]}\,d\lambda \quad (30)$$

Hence the effect of misses is equivalent to adding white noise to the spectral density of a sampled-data system operating without misses.

Spectral Densities of a Markov Signal

Assume the spectral density of a signal before sampling is:

$$G_{s_k s_k}(s)=\frac{2\beta\sigma^2}{-s^2+\beta^2}$$

Then:

$$G_{s_k s_k}{}^*(s)=\frac{\sigma^2\bar{T}_k[1-M_k(s+\beta)M_k(-s+\beta)]}{[1-M_k(s+\beta)][1-M_k(-s+\beta)]} \quad (31)$$

If $\beta\to\infty$ and σ^2 is bounded, then the sampled Markov signal will approach a white spectral density:

$$G_{s_k s_k}{}^*=\sigma^2\bar{T}_k \quad (32)$$

It should be noted that the continuous spectral density before sampling for this limiting case tends to zero.

TRANSFORMATION OF INTEGRAL EQUATIONS

The integral equations which must be solved are shown to be:

$$\sum_{k=1}^{n}\int_0^\infty W_{jk}(\tau_1)\phi_{i_k'i_k}{}^*(\tau_2-\tau_1)d\tau_1 = \sum_{k=1}^{n}\int_{-\infty}^{\infty} D_{jk}(\tau_1)\phi_{i_k's_k}(\tau_2-\tau_1)d\tau_1$$
$$\text{for } \tau_2\geqq 0,\quad k'=1, 2, \ldots, n$$

These equations should be modified before transformation can be applied. Let $f_{jk'}(\tau)$ be defined as:

$$f_{jk'}(\tau)=0 \qquad \text{for } \tau\geqq 0 \quad (33)$$

Then the integral equations can be rewritten as:

$$\sum_{k=1}^{n}\int_0^\infty W_{jk}(\tau_1)\phi_{i_k'i_k}{}^*(\tau_2-\tau_1)d\tau_1 = \sum_{k=1}^{n}\int_{-\infty}^{\infty} D_{jk}(\tau_1)\phi_{i_k's_k}(\tau_2-\tau_1)d\tau_1+f_{jk'}{}'(\tau_2)$$
$$\text{for all values of } \tau_2 \text{ and } k'=1, 2, \ldots, n \quad (34)$$

With a stable system and the property of the correlation function $\phi(\tau)$ or $\phi^*(\tau)$ for large τ(i.e., $\phi(\tau)\to 0$ as $|\tau|\to\infty$), it is quite obvious that $f_{jk'}(\tau)$ will also approach zero for very large negative τ. In general, $f_{jk'}(\tau)$ will not have any delta functions.

The application of the two-sided Laplace transform in solving equation 34 depends on the existence of an analytic strip common to all the functions being transformed. For a physically realizable and stable system, the transform of the weighting functions, which are defined as transfer functions of the system, can only have singularities in the left half plane. The transform of $\phi_{i_k'i_k}{}^*(\tau)$ or $\phi_{i_k's_k}(\tau)$ will have an analytic strip around the imaginary axis. Its singularities in the left half plane arise from the part of the correlation function with positive argument and those in the right half plane, from the part of the correlation function with negative argument. The transform of the arbitrarily defined function $f_{jk'}(\tau)$ can evidently possess singularities in the right half of the plane only. Thus with an appropriate specification of the desired transfer functions, the existence of a common analytic strip is always obtainable. Thus the inverse transform is unique. It is noticed that this analytic strip contains the imaginary axis which can be chosen as the path for inverse transformation. This line integration can always be replaced by contour integration. Closing the contour in the left half plane will give the function for positive t.

Let the two-sided Laplace transform be taken on both sides of equation 34.

$$\sum_{k=1}^{n} G_{i_k'i_k}{}^*(s)\,Y_{jk}(s) = \sum_{k=1}^{n}(Y_d)_{jk}(s)G_{i_k's_k}(s)+F_{jk'}{}^-(s)$$
$$k'=1, 2, \ldots, n \quad (35)$$

where

$$Y_{jk}(s)=\int_0^\infty W_{jk}(t)e^{-st}dt$$
(System transfer functions) (36)

$$(Y_d)_{jk}(s)=\int_{-\infty}^{\infty} D_{jk}(t)e^{-st}dt$$
(Ideal transfer functions) (37)

$$G_{i_k'i_k}{}^*(s)=\int_{-\infty}^{\infty}\phi_{i_k'i_k}{}^*(t)e^{-st}dt = G_{i_k'i_k} \text{ for } k\neq k'$$
(Sampled power spectral densities of inputs) (38)

$$G_{i_k's_k}(s)=\int_{-\infty}^{\infty}\phi_{i_k's_k}(t)e^{-st}dt$$
(Continuous cross-power spectral densities of inputs and signals) (39)

$$F_{jk'}{}^-(s)=\int_{-\infty}^{0} f_{jk'}(t)e^{-st}dt \quad (40)$$

Define

$$N_{jk'}(s)=\sum_{k=1}^{n}(Y_d)_{jk}(s)G_{i_k's_k}(s) \quad (4$$

Then equation 35 can be written as:

$$\sum_{k=1}^{n} G_{i_k'i_k}{}^*(s)\,Y_{jk}(s)=N_{jk'}(s)+F_{jk'}{}^-(s)$$
$$k'=1, 2, \ldots, n \quad (4$$

In matrix form:

$$\mathbf{GY}_j=\mathbf{N}_j+\mathbf{F}_j{}^- \text{ for } j=1, 2, \ldots, m \quad (4$$

where the square matrix $\mathbf{G}$ is identi for all output terminals and is not s gular, and the vector $\mathbf{F}_j{}^-$ which is s unknown at this moment can only ha singularities in the right half plane. W $s=jw$, $\mathbf{G}$ is Hermitian. Inverting equ tion 43, transfer function vectors are:

$$\mathbf{Y}_j=\mathbf{G}^{-1}(\mathbf{N}_j+\mathbf{F}_j{}^-)=\frac{\mathbf{A}}{|G|}(\mathbf{N}_j+\mathbf{F}_j{}^-) \quad (4$$

In this equation, $|G|$ is the determinant $\mathbf{G}$ and should not have any root on t imaginary axis. $\mathbf{A}$ is the adjoint of $\mathbf{G}$.

In complex domain, the minimum mea square errors are

$$\overline{(\epsilon_j{}^2)}_{\min}=\sum_{k'=1}^{n}\sum_{k=1}^{n}\frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}\times [(Y_d)_{jk'}{}^-(Y_d)_{jk}G_{s_k's_k}- Y_{jk'}{}^-Y_{jk}G_{i_k'i_k}{}^*]ds$$
$$=\sum_{k'=1}^{n}\sum_{k=1}^{n}\frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}\times [(Y_d)_{jk'}{}^-(Y_d)_{jk}G_{s_k's_k}- Y_{jk'}{}^-(Y_d)_{jk}G_{i_k's_k}]ds \quad (4$$
$$j=1, 2, \ldots, m$$

Here the integration path lies in the an lytic strip common to all these tra formed functions and $Y^-(s)\triangleq Y(-s)$.

Solution of Transfer Functions the Optimum Filters

FACTORIZATION OF SPECTRAL DENSI DETERMINANT

The feasibility of using transfo methods to solve the integral equatio depends upon the factorization of spectral-density determinant $|G|$ into t functions $G^+(s)G^-(s)$, where $G^+(s)$ l poles and zeros in the left half plane a $G^-(s)$ has poles and zeros in the right h plane. Theoretically, this can alwa be done if the Paley-Weiner criterion satisfied. Practically, however, the g eral procedure specified by Wiener is v difficult computationally unless the fu tion is of certain convenient forms.

It is usually assumed that the spect densities of continuous inputs can be

sed as rational functions in s. Thus off-diagonal elements of $|G|$ are all nal functions in s. However, the ents along the diagonal are the pled-power spectral densities. They not be rational functions in s. This es additional difficulty in effecting the rization, which is not encountered in simple two-pole systems.

wo important classes of sampling will considered. In both cases the spec- densities before sampling are all med to be rational functions in s. he first class, the sampling density rating function is rational in s. will include, for example, those sam- g intervals governed by the gamma ity function.[2] Thus the spectral- ity determinant, in this case, will be even function of s and the factoriza- can easily be obtained.

the second class, the sampling inter- density is discrete. This will consist ne case of missed samples and the case re a finite number of possible sampling rvals exists. In either case, the spec- density after sampling is a function ^{sT_n}, thus the determinant $|G|$ will be a ction in both s and e^{sT_n}. There is no way to carry out the factorization of function. In order to overcome this culty, the approximation of e^{sT_n} by onal function should be used. The é approximants (see the Appendix) particularly useful in this case. The ner the degrees of numerator and ominator polynomials of the approxi- t, the better the approximation. wever, the complexity of the system gned will be increased accordingly. s approach is justified by the fact that signals usually have power concen- ed in low frequencies and the syn- sized system is low pass. It should be ed that it is usually desirable to choose polynomials of the same degree for the nerator and denominator of the ap- ximant in order to obtain the simplest n for the approximated sampled ctral density.

Thus it has been shown that for these classes of sampling the power spec- density determinants will be even ctions in s. The synthesis procedures then be similar to those for continu- inputs.[4]

IMUM FILTERS FOR TWO IMPORTANT LASSES OF SAMPLING

he previous section has demonstrated fact that, for two classes of sampling, sampled spectral densities are rational ctions in s or can be approximated rational functions in s. Thus the ctral-density determinant can be easily mposed such that:

$$|G| = G^+(s)G^-(s)$$

where all the poles and zeros of $G^+(s)$ must lie in the left half plane and $G^-(s)$, which is equal to $G^+(-s)$, can only have poles and zeros in the right half plane. There should not be any finite zero along the imaginary axis.

The standard procedure in finding the physically realizable transfer functions for this Wiener type of problem can now be adopted. Thus:

$$\mathbf{Y}_j = \frac{\mathbf{A}}{G^+G^-}\,[\mathbf{N}_j{}^+\mathbf{F}_j{}^-]$$

or

$$G^+\mathbf{Y}_j = \frac{1}{G^-}\mathbf{A}\mathbf{N}_j + \frac{1}{G^-}\mathbf{A}\mathbf{F}_j{}^- \tag{46}$$

The matrix on the left side of the equation can have poles only in the left half plane. Hence it follows that

$$G^+\mathbf{Y}_j = \left\{\frac{1}{G^-}\mathbf{A}\mathbf{N}_j\right\}^+ + \left\{\frac{1}{G^-}\mathbf{A}\mathbf{F}_j{}^-\right\}^+ \tag{47}$$

where $\{\ \}^+$ denotes that part of the function which has poles in the left half plane only. The elements in $(1/G^-)\mathbf{A}\mathbf{N}_j$ are all known. Therefore:

$$\left\{\frac{1}{G^-}\mathbf{A}\mathbf{N}_j\right\}^+ = \int_0^\infty \mathbf{g}(t)e^{-st}dt \tag{48}$$

where

$$\mathbf{g}(t) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}\frac{1}{G^-}\mathbf{A}\mathbf{N}_j e^{ts}ds$$

The left half plane poles in $(1/G^-)\mathbf{A}\mathbf{F}_j{}^-$ can only come from the matrix $\mathbf{A}$. Since $\mathbf{F}_j{}^-$ is still unknown, the residues or coefficients associated with these poles can not be determined yet. Thus

$$\left\{\frac{1}{G^-}\mathbf{A}\mathbf{F}_j{}^-\right\}^+$$

will be a column vector whose kth element contains poles which are the left half plane poles from the kth row elements of matrix $\mathbf{A}$ and can be expressed:

$$\sum_{i,m_i}\frac{C_{im_i}}{(s-\gamma_i)^{m_i}}$$

where γ_i are the left half plane poles, which can be multiple, and C_{im_i} are the unknown coefficients. A typical transfer function can now be expressed by

$$Y_{jk} = \frac{1}{G^+}\left[\left\{\frac{1}{G^-}\sum_{k'=1}^{n}A_{k'k}N_{jk'}\right\}^+ + \sum_{i,m_i}\frac{C_{im_i}}{(s-\gamma_i)^{m_i}}\right] \tag{49}$$

In order to determine these coefficients, Y_{jk} are substituted into the original system equations.

It must be true then:

$$\{\mathbf{G}\mathbf{Y}_j\}^+ = \{\mathbf{N}_j\}^+ \tag{50}$$

or, more explicitly:

$$\left\{\sum_{k=1}^{n}G_{i_k'i_k}{}^*(s)\,Y_{jk}(s)\right\}^+ = \left\{\sum_{k=1}^{n}(Y_d)_{jk}(s)G_{i_k's_k}(s)\right\}^+$$

for $k' = 1, 2, \ldots, m$

Hence a set of independent linear algebraic equations in terms of C_{im_i} will be obtained. These coefficients can thus be uniquely determined.

In working an actual problem, it is more desirable to modify the expression for Y_{jk} as given by equation 50 in order to minimize the accumulated computational errors. Since the left half plane poles of the cofactors $A_{kk'}$ will eventually be cancelled by the poles of G^+, the transfer functions will take the form of equation 51:

$$Y_{jk}(s) = \frac{\left[\prod_u(s-a_u)\right]\left[C_n s^n + C_{n-1}s^{n-1} + \ldots + C_0\right]}{\left[\prod_v(s-b_v)\right]\left[\prod_l(s-d_l)\right]} \tag{51}$$

where

b_v = the zeros of G^+

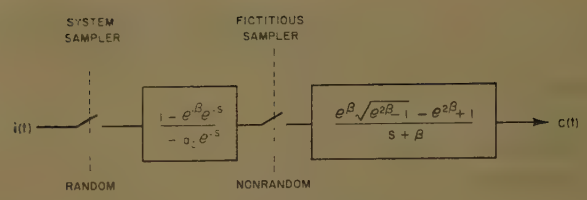
d_l = the left half plane poles of the ideal transfer functions vector $(\mathbf{Y}_d)_j$ if they are different from those of the spectral densities or have not been cancelled by the poles of G^+

a_u = the poles of G^+, which have not been cancelled by the poles of the kth row element of

$$\left\{\frac{1}{G^-}\mathbf{A}(\mathbf{N}_j{}^+\mathbf{F}_j{}^-)\right\}^+$$

vector, and where $\{C_n\}$ are the new unknown coefficients which will be a linear combination of the old coefficients $\{C_{im_i}\}$. This set of new coefficients is always greater in number than the old ones and, therefore, more simultaneous equations have to be solved. However, with the aid of a digital computer, the solution for these equations can easily be obtained. As a rule, only $(n-1)$ equations of the n equations associated with the same pole b_v of the transfer functions are independent. This procedure will be amplified by referring to example 2, discussed later. The poles of Y_{jk} are usually simple; however, they can be multiple.

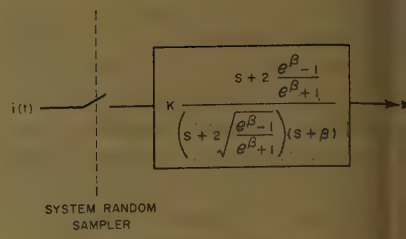
The synthesis procedure for determining the optimum transfer function vector $\mathbf{Y}_j$ will then be as follows:



A—Exact system

Fig. 4 (left and right). Example 1; probability density function is:

$$\sum_{n=1}^{\infty} 0.8(0.2)^{n-1}\delta(T-n)$$



B—Approximate system where

$$K=\frac{1+e^{-\beta}}{2}\,\frac{2\,\frac{e^{\beta}-1}{e^{\beta}+1}+\beta}{2\sqrt{\frac{e^{\beta}-1}{e^{\beta}+1}}+\beta}$$

1. When the spectral densities after sampling are rational functions in s, the spectral-density determinant $|G|$ can be factored into $G^{+}(s)$ and $G^{-}(s)$ directly. When the spectral densities after sampling are functions of e^{sTn}, they should be approximated by rational function in s before factorization.

2. Express the system transfer functions with the form which is given in equation 51.

3. Determine the unknown coefficients by substituting $\mathbf{Y}_j$ into the original system equations.

It is evident that the poles of the transfer functions will consist of two parts. First, all the system transfer functions will have poles which are the zeros of $G^{+}(s)$. These poles, in general, are completely different from the left half plane poles of all the spectral densities. Second, the transfer functions associated with a particular output terminal may contain the left half plane poles from the desired transfer functions vector $(\mathbf{Y}_d)_j$.

Illustrative Examples

Example 1. The validity of using the fractional function approximation of e^{sTn} to the optimum synthesis of system will now be shown. A two-pole system will be used so that the mean-square error of the system from the approximate solution can be compared with that from the exact solution.

Let us consider the design of an optimum filter with a nominal sampling period of 1 second and the probability of a miss of 0.2. Assume the signal is Markov with the form

$$G_{ss}(s)=\frac{2\beta}{-s^2+\beta^2}$$

The noise is white but with a finite mean-square value of 0.6. The signal and noise are uncorrelated. The desired operation is pure filtering.

Thus

$$Y_d(s)=1$$

In this simple problem, the equation to be solved is:

$$G_{ii}^{*}(s)\,Y(s)=Y_d(s)G_{ss}(s)+F^{-}(s) \qquad (52)$$

The sampled spectral density can be found by using equations 29 and 32.

$$G_{ii}^{*}(s)=G_{ss}^{*}(s)+G_{NN}^{*}(s)$$

$$=\left[\frac{1-e^{-2\beta}}{(1-e^{-\beta}e^{-s})(1-e^{-\beta}e^{s})}+\frac{0.2}{0.8}\times 1\right]+\frac{1}{0.8}\times 0.6$$

$$=e^{-\beta}\,\frac{2e^{\beta}-e^{s}-e^{-s}}{(1-e^{-\beta}e^{-s})(1-e^{-\beta}e^{s})} \qquad (53)$$

The numerator can be factored as:

$$2e^{\beta}-e^{s}-e^{-s}=\frac{1}{a_0}\,(1-a_0e^{-s})(1-a_0e^{-s})$$

where

$$a_0=e^{\beta}-\sqrt{e^{2\beta}-1}$$

Thus

$$G^{+}(s)=b\,\frac{1-a_0e^{-s}}{1-e^{-\beta}e^{-s}}$$

with

$$b=\sqrt{\frac{e^{-\beta}}{e^{\beta}-\sqrt{e^{2\beta}-1}}}$$

The exact solution for the optimum transfer function can then be found.

$$Y_0(s)=\frac{1}{G^{+}}\left\{\frac{Y_dG_{ss}}{G^{-}}\right\}^{+}$$

$$=\frac{1-e^{-\beta}e^{-s}}{1-a_0e^{-s}}\;\frac{e^{\beta}\sqrt{e^{2\beta}-1}-e^{2\beta}+1}{s+\beta} \qquad (54)$$

The exact system is shown in Fig. 4(A). The minimum mean-square error is:

$$\overline{\epsilon_0^2}=1-\frac{(e^{\beta}-e^{-\beta})(e^{\beta}-\sqrt{e^{2\beta}-1})}{2\beta} \qquad (55)$$

Let the approximants chosen be as follows:

$$e^{-s}\simeq\frac{1-\frac{1}{2}s}{1+\frac{1}{2}s}$$

and

$$e^{s}\simeq\frac{1+\frac{1}{2}s}{1-\frac{1}{2}s}$$

Then

$$(G_{ii})_1^{*}(s)=\frac{2e^{\beta}}{e^{\beta}+1}\times\frac{\left(2\sqrt{\frac{e^{\beta}-1}{e^{\beta}+1}}+s\right)\left(2\sqrt{\frac{e^{\beta}-1}{e^{\beta}+1}}-s\right)}{\left(2\,\frac{e^{\beta}-1}{e^{\beta}+1}+s\right)\left(2\,\frac{e^{\beta}-1}{e^{\beta}+1}-s\right)}$$

$$Y_1(s)=\frac{1+e^{-\beta}}{2}\times\frac{\left(2\,\frac{e^{\beta}-1}{e^{\beta}+1}+\beta\right)\left(s+2\,\frac{e^{\beta}-1}{e^{\beta}+1}\right)}{\left(2\sqrt{\frac{e^{\beta}-1}{e^{\beta}+1}}+\beta\right)\left(s+2\sqrt{\frac{e^{\beta}-1}{e^{\beta}+1}}\right)(s$$

This approximate system is shown Fig. 4(B). Due to the approxima used, $Y_1(s)$ does not satisfy equatio exactly.

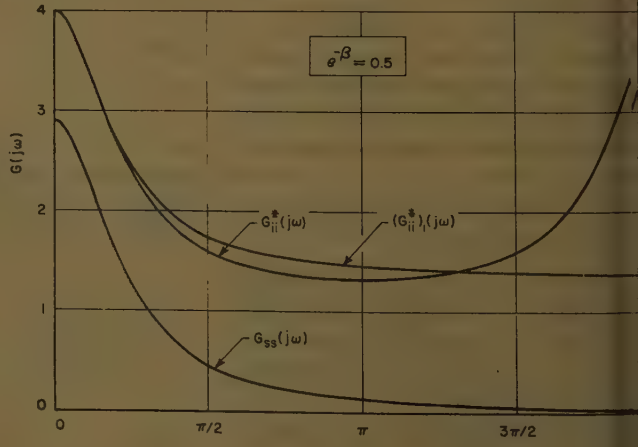


Fig. 5. Power spectral densities in example 1

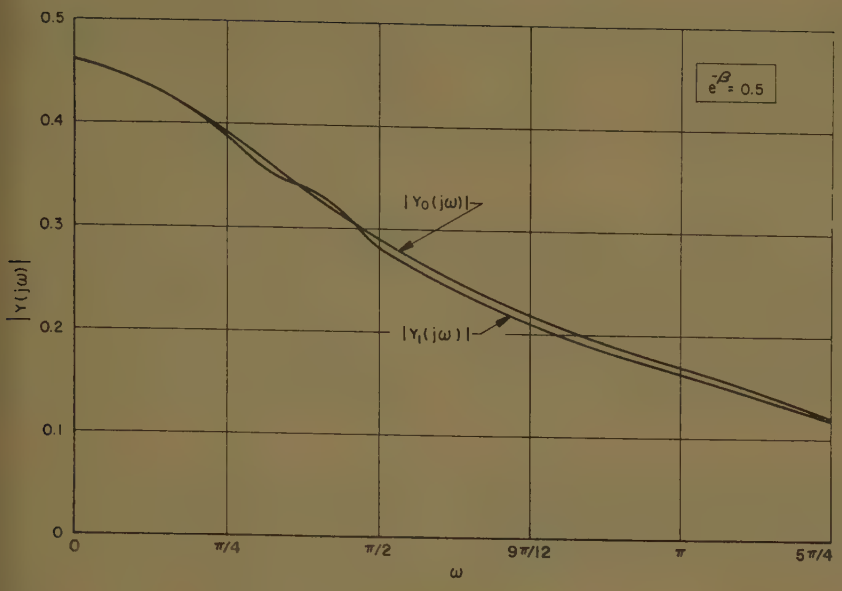


Fig. 6. Gain diagrams in example 1

$$\overline{\epsilon_1^2}=\frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}[Y(-s)Y(s)G_{ii}^*(s)-2Y(-s)Y_d(s)G_{is}(s)+Y_d(-s)Y_dG_{ss}(s)]ds \quad (58)$$

Hence, by using contour integration:

$$\overline{\epsilon_1^2}=\left[\frac{(1+e^{-\beta})(2a+\beta)}{2(2\sqrt{a}+\beta)}\right]^2\left\{\frac{4a^2+2\beta-\beta^2}{2\beta(4a-\beta^2)}-\frac{(4a^2-\beta^2)(10\beta^2-8a)}{8\beta^3(4a-\beta^2)^2}+\frac{(1+e^{-2\beta})(4a^2-\beta^2)}{4\beta(1-e^{-2\beta})(4a-\beta^2)}+\sum_{n=1}^{\infty}\times\frac{n\pi(4a^2-\beta^2+4n^2\pi^2)(4a-\beta^2+4n^2\pi^2)+16n^2\pi^2-4(a^2-a)}{2n\pi[(4a-\beta^2+4n^2\pi^2)^2+16n^2\pi^2]\times(n^2\pi^2+1)}+\frac{(a^2-a)(2-e^{-(\beta+2\sqrt{a})}-e^{2\sqrt{a}-\beta})}{\sqrt{a}(\beta^2-4a)(1-e^{2\sqrt{a}-\beta})(1-e^{-(\beta+2\sqrt{a})})}\right\}-\frac{1+e^{-\beta}}{2\beta}\left(\frac{2a+\beta}{2\sqrt{a}+\beta}\right)^2+1 \quad (59)$$

where

$$\triangleq\frac{e^{\beta}-1}{e^{\beta}+1}$$

The infinite series behaves as

$$\sum_{n=1}^{\infty}\frac{1}{2(n^2\pi^2+1)}$$

and, therefore, is convergent.

Fig. 5 shows the spectral densities for real frequencies. It is seen that the approximated sampled spectral density gives a reasonably good fit to the exact one up to frequency $4\pi/3$. After that, it deviates from the exact one considerably. However, the magnitude of the signal spectral density $G_{ss}(jw)$ at this frequency range is very small. The filter is certainly of low pass (Fig. 6). Thus, the error introduced by the approximation is expected to be negligible.

Fig. 7 shows the variation of

$$\overline{\epsilon_0^2}$$

and

$$\overline{\epsilon_1^2}$$

as a function of β. For small β, which implies narrow bandwidth of the spectral densities, these two errors agree very closely. They deviate very slightly as β becomes greater than 1. It can easily be shown that both

$$\overline{\epsilon_0^2}$$

and

$$\overline{\epsilon_1^2}$$

approach zero as $\beta\to 0$ and approach 1 as $\beta\to\infty$.

Example 2. The synthesis procedure for a multipole filter will now be illustrated. For simplicity, a system with two inputs and one output is chosen. The signal portions of the inputs are assumed to be the same. However, they are corrupted by different noises. These noises are not correlated with the signals and are assumed to be white with finite mean-square values. Thus the correlation functions are:

$$\phi_{S_1S_1}(\tau)=\phi_{S_2S_2}(\tau)=e^{-|\tau|}$$
$$\phi_{N_1N_1}(\tau)=0.5 \quad \text{for } \tau=0$$
$$=0 \quad \text{otherwise}$$

and

$$\phi_{N_2N_2}(\tau)=0.2 \quad \text{for } \tau=0$$
$$=0 \quad \text{otherwise}$$

The samplers have a nominal sampling period of 1 second but with missing probabilities of 0.1 and 0.2 respectively. The function of the filter is to extract the signal from these two different channels; see Fig. 8.

The power spectral densities for signals before sampling are:

$$G_{S_1S_1}(s)=G_{S_2S_2}(s)=\frac{2}{-s^2+1}$$

and

$$G_{S_1S_2}(s)=G_{S_2S_1}(s)=\frac{2}{-s^2+1}$$

Thus

$$G_{i_1S_1}(s)=G_{i_1S_2}(s)=G_{i_2S_1}(s)=G_{i_2S_2}(s)=\frac{2}{-s^2+1}$$

and

$$G_{i_1i_2}(s)=G_{i_2i_1}(s)=\frac{2}{-s^2+1}$$

The spectral densities for inputs after sampling are:

$$G_{i_1i_1}^*(s)=\frac{1.622-0.245(e^{-s}+e^{s})}{1.135-0.368(e^{-s}+e^{s})} \quad (60)$$

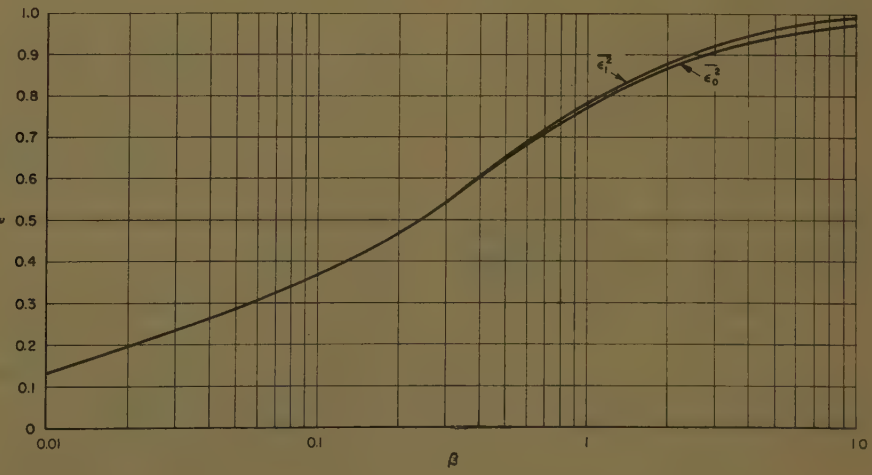


Fig. 7. Minimum mean-square errors in example 1

$$G_{i_2i_2}^*(s)=\frac{1.433-0.184(e^{-s}+e^{s})}{1.135-0.368(e^{-s}+e^{s})} \quad (61)$$

The desired transfer functions are:

$$(Y_d)_{11}(s)=(Y_d)_{12}(s)=\frac{1}{2}$$

The equations to be solved are:

$$G_{i_1i_1}^*Y_{11}+G_{i_1i_2}Y_{12}=(Y_d)_{11}G_{i_1s_1}+(Y_d)_{12}G_{i_1s_2}+F_{11}^-$$

$$G_{i_2i_1}Y_{11}+G_{i_2i_2}^*Y_{12}=(Y_d)_{11}G_{i_2s_1}+(Y_d)_{12}G_{i_2s_2}+F_{12}^- \quad (62)$$

Now let

$$e^{-s}\simeq\frac{1-\frac{1}{2}s}{1+\frac{1}{2}s}$$

and

$$e^{s}\simeq\frac{1+\frac{1}{2}s}{1-\frac{1}{2}s}$$

Then

$$G_{i_1i_1}^*(s)\simeq 1.129\,\frac{-s^2+2.144}{-s^2+0.853} \quad (63)$$

$$G_{i_2i_2}^*(s)\simeq 0.962\,\frac{-s^2+2.367}{-s^2+0.853} \quad (64)$$

The spectral-density matrix is

$$\mathbf{G}=\begin{bmatrix} 1.129\,\dfrac{-s^2+2.144}{-s^2+0.853} & \dfrac{2}{1-s^2} \\ \dfrac{2}{1-s^2} & 0.962\,\dfrac{-s^2+2.367}{-s^2+0.853} \end{bmatrix} \quad (65)$$

Hence its determinant can be factored into two functions defined as:

$$G^+=1.087\,\frac{(s+0.9558)(s+2.0649)(s+0.8518+j0.2437)(s+0.8518-j0.2437)}{(s+0.923)^2(s+1)^2}$$

$$G^-=\frac{(s-0.9558)(s-2.0649)(s-0.8518+j0.2437)(s-0.8518-j0.2437)}{(s-0.923)^2(s-1)^2}$$

The matrices $\mathbf{N}_1$ and $\mathbf{A}$ are:

$$\mathbf{N}_1=\begin{bmatrix} \dfrac{2}{-s^2+1} \\ \dfrac{2}{-s^2+1} \end{bmatrix} \quad (66)$$

$$\mathbf{A}=\begin{bmatrix} 0.962\,\dfrac{-s^2+2.367}{-s^2+0.853} & \dfrac{-2}{-s^2+1} \\ \dfrac{-2}{-s^2+1} & 1.129\,\dfrac{-s^2+2.144}{-s^2+0.853} \end{bmatrix} \quad (67)$$

By using equation 47:

$$G^+\mathbf{Y}_1=\left\{\frac{1}{G^-}\mathbf{AN}_1\right\}^+ +\left\{\frac{1}{G^-}\mathbf{AF}_1^-\right\}^+=\begin{bmatrix} \dfrac{-6.802s-7.503}{(s+1)^2}+\dfrac{7.373}{s+0.923}+\dfrac{A_{11}}{s+0.923}+\dfrac{A_{12}}{s+1} \\ \dfrac{-6.687s-7.388}{(s+1)^2}+\dfrac{7.373}{s+0.923}+\dfrac{A_{21}}{s+1}+\dfrac{A_{22}}{s+0.923} \end{bmatrix} \quad (68)$$

It is obvious that the right side matrix can be written as:

$$\begin{bmatrix} \dfrac{a_2s^2+a_1s+a_0}{(s+1)^2(s+0.923)} \\ \dfrac{b_2s^2+b_1s+b_0}{(s+1)^2(s+0.923)} \end{bmatrix}$$

where a_2, a_1, and a_0 are the new unknown coefficients which are linear combinations of A_{11} and A_{12}; and b_2, b_1, and b_0 are another set of new unknown coefficients which are linear combinations of A_{21} and A_{22}. Thus the transfer function vector is:

$$\mathbf{Y}_1=\begin{bmatrix} \dfrac{(s+0.923)(a_2s^2+a_1s+a_0)}{(s+0.9558)(s+2.0649)(s+0.8518+j0.2437)(s+0.8518-j0.2437)} \\ \dfrac{(s+0.923)(b_2s^2+b_1s+b_0)}{(s+0.9558)(s+2.0649)(s+0.8518+j0.2437)(s+0.8518-j0.2437)} \end{bmatrix} \quad (69)$$

Each element is of the form expressed in equation 51. It is observed that quick examination on the right side of equation 68 will yield equation 69 immediately.

By substituting equation 69 into equation 62, six independent equations for these six coefficients will be obtained from evaluating the residues of the left half plane poles. These poles are the four poles of the transfer functions and the one at $s=-1$. It should be noted that the two equations obtained by using the same transfer-function's pole in the set of optimum equations 62 will be dependent. However, the two equations associated with the pole at $s=-1$ are independent. Thus the optimum transfer functions are (equations 70 and 71):

$$Y_{11}(s)=\frac{(s+0.923)(0.5882s^2+1.0065s+0.468)}{(s+0.9558)(s+2.0649)(s+0.8518+j0.2437)(s+0.8518-j0.2437)} \quad (70)$$

$$Y_{12}(s)=\frac{(s+0.923)(0.6744s^2+1.1487s+0.527)}{(s+0.9558)(s+2.0649)(s+0.8518+j0.2437)(s+0.8518-j0.2437)} \quad (71)$$

The minimum mean-square error is:

$$\overline{\epsilon_{\min}^2}=0.5942$$

If the input of the first channel alone is used, the optimum filter and its minimum mean-square error are:

$$Y_a(s)=\frac{1-0.368e^{-s}}{1-0.155e^{-s}}\,\frac{0.577}{s+1};\quad \overline{(\epsilon_a^2)_{\min}}=0.7864$$

On the other hand, if the input of the second channel only is used, they are:

$$Y_b(s)=\frac{1-0.368e^{-s}}{1-0.131e^{-s}}\,\frac{0.6447}{s+1};\quad \overline{(\epsilon_b^2)_{\min}}=0.7072$$

It is clear then that the multipole filter has better performance than the corresponding 2-pole filters at the expense of a more complicated system.

If equation 68 is chosen for the expression of the transfer functions, then the four independent equations for the unknown coefficients can be obtained by evaluating the residues associated with the four poles of the transfer functions either from the first or second equation of the simultaneous equations 62. It should be true, theoretically, that the residues associated with the pole $s=-1$ on the left sides of these two system equations are 1 and, therefore, are independent of the unknown coefficients. However, because of the computational errors, unfortunately these residues will not be exactly equal to the expected values. On the other hand, the second expressions usually tend to minimize the computational errors which will inevitably occur.

Conclusions

This paper has shown that with stationary inputs and the assumption of statistically independent sampling processes, an optimal multipole filter in the Wiener sense can be synthesized.

1. The cross-power spectral densities of the sampled inputs and cross-power spectral densities of the sampled and the co

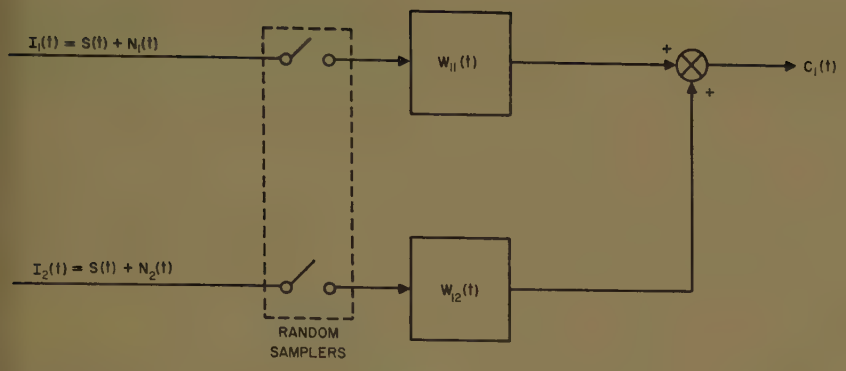


Fig. 8. Example 2, 2+1 pole filter

inuous inputs are unaffected by the ampling processes. The auto-power spec-ral densities of the sampled inputs are, on he other hand, influenced by the sampling rocesses.

2. If the sampled spectral densities are ot rational functions in s, they have to e approximated, at least for small s, by his kind of function. Since the signals sually have power concentrated in the ow frequencies, the error introduced by his approximation is negligible.

3. The set of generalized Wiener-Hopf equations is solved by transform method. Factorization of the spectral-density determinant and the method of undetermined coefficients are employed in obtaining the optimal transfer functions, which are rational functions in s and therefore can easily be synthesized with a finite number of lumped linear passive components.

Appendix. Padé Approximants for e^{-s}

$$\frac{1-s}{1} \qquad \frac{1-s+\frac{s^2}{2!}}{1}$$

$$\frac{1-\frac{1}{2}s}{1+\frac{1}{2}s} \qquad \frac{1-\frac{2}{3}s+\frac{1}{3}\frac{s^2}{2!}}{1+\frac{1}{3}s}$$

$$\frac{1-\frac{1}{3}s}{1+\frac{2}{3}s+\frac{1}{3}\frac{s^2}{2!}} \qquad \frac{1-\frac{1}{2}s+\frac{1}{6}\frac{s^2}{2!}}{1+\frac{1}{2}s+\frac{1}{6}\frac{s^2}{2!}}$$

References

1. Analysis and Synthesis of Linear Systems Operating on Randomly Sampled Data, R. E. Kalman. *Doctoral Dissertation*, Department of Electrical Engineering, Columbia University, New York, N. Y., 1957.

2. The Synthesis of Optimum Random Sampling Systems, A. R. Bergen. *Technical Report T-T/133*, Electronics Research Laboratories, Columbia University, 1957.

3. Extrapolation, Interpolation and Smoothing of Stationary Time Series (book), Norbert Wiener. John Wiley & Sons, Inc., New York, N. Y., 1949.

4. On the Optimum Synthesis of Multipole Control Systems in the Wiener Sense, H. C. Hsieh, C. T. Leondes. *Transactions*, Professional Group on Automatic Control, Institute of Radio Engineers, New York, N. Y., vol. AC-4, no. 2, Nov. 1959, pp. 16–29.

5. On the Optimum Synthesis of Sampled Data Multipole Filters with Random and Nonrandom Inputs, H. C. Hsieh, C. T. Leondes. *Ibid.*, vol. AC-5, no. 3, Aug. 1960, pp. 193–208.

Sampled-Data Control Systems with Transport Lag by Mitrović's Algebraic Method

DRAGOSLAV ŠILJAK
NONMEMBER AIEE

TRANSPORT LAG in sampled-data control systems influences their stability. If this influence is studied as the function of transport lag and other system parameters, the design procedure becomes cumbersome. This paper, which applies Mitrović's algebraic method,[1–3] presents a new way to analyze and synthesize various sampled-data control systems such as the cyclic variable-rate, multirate, and those with nonsynchronized samplers. Other related problems are greatly facilitated by this method also.

Paper 61-743, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Ithaca, N. Y., June 18–23, 1961. Manuscript submitted February 21, 1961; made available for printing April 25, 1961.

Dragoslav Šiljak is with the University of Belgrade, Belgrade, Yugoslavia.

Making use of flat-top-pulse or trapezoid-pulse approximations,[4] the proposed design procedure can be extended to cover sampled-data control systems with finite sampling duration. Moreover, this technique can deal effectively with problems appearing in sampled-data systems in which hidden instability[5,6] occurs.

A simple single-loop system with transport lag T_t and the conventional linear part characterized by a transfer function $G(s)$, shown in Fig. 1, will be considered representative in the analysis. Procedure and conclusions outlined herein are applicable, with only slight modifications, to other types mentioned.

Before analyzing the system, a necessary preliminary is to determine the z-transfer function $G(z, T_t)$, corresponding to $e^{-T_t s}G(s)$. Since the method to be used is based upon study of the system-characteristic equation having the algebraic form

$$a_n z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0 = 0 \qquad (1)$$

$G(z, T_t)$ should be a rational function in z. T_t is not generally a multiple integer of the sampling period T. Hence, to obtain $G(z, T_t)$ in the required form, T_t must be expressed as

$$T_t = kT - mT \qquad (2)$$

where k is the upper integer of T_t/T and m is a fraction $(0 < m \leq 1)$. Finding the modified z-transform $G(z, m)$ corresponding to $G(s)$ and substituting m with $k - T_t/T$ from equation 2, the required form of function $G(z, T_t)$ is

$$G(z, T_t) = z^{-(k-1)}G(z, m)\,|_{m=k-T_t/T} = z^{-(k-1)}G(z, k - T_t/T) \qquad (3)$$

Using equation 3, system behavior is determined by an algebraic equation of form 1, and Mitrović's algebraic method is applied. For convenience, this method is stated briefly.

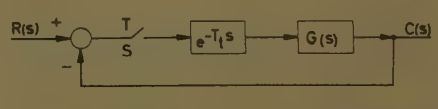


Fig. 1. Block diagram of sampled-data control system with transport lag

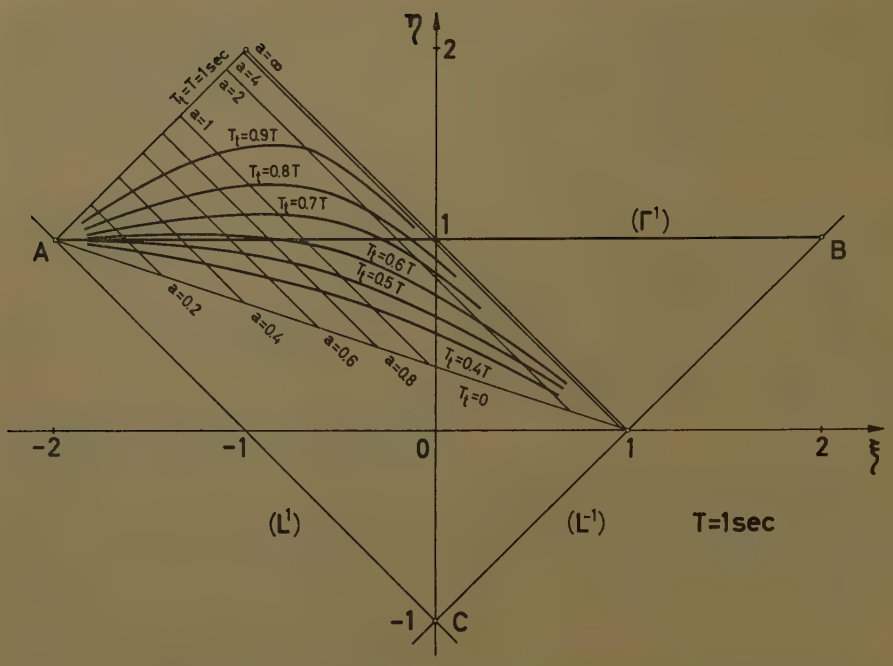


Fig. 2. Correlation between system parameters a and T_t

Considering coefficients a_1 and a_0 in equation 1 as variables ξ and η, and using relationships

$$z=-\omega_z\zeta_z+j\omega_z\sqrt{1-\zeta_z^2}=e^{sT}$$
$$=e^{(-\omega_n\zeta+j\omega_n\sqrt{1-\zeta^2})} \quad (4)$$

where ω_n is undamped natural frequency and ζ is the relative damping coefficient, will get

$$\xi=a_2\phi_2(-\cos\omega_nT\sqrt{1-\zeta^2})e^{-\omega_n\zeta T}+$$
$$a_3\phi_3(-\cos\omega_nT\sqrt{1-\zeta^2})e^{-2\omega_n\zeta T}+\ldots$$
$$\eta=-e^{-2\omega_n\zeta T}[a_2\phi_1(-\cos\omega_nT\sqrt{1-\zeta^2})+$$
$$a_3\phi_2(-\cos\omega_nT\sqrt{1-\zeta^2})e^{-\omega_n\zeta T}+\ldots] \quad (5)$$

These equations represent the loci of points in the $0\xi\eta$-plane corresponding to roots with settling time, damping coefficient, or undamped natural frequency being constant, depending on which variable among them is considered as constant. Plotting of these characteristic curves offers no particular difficulty since the table of functions e^{-x} and Table I in reference 3 of the functions $\phi_i(-\cos x)$ may be used.

The essential point after plotting these curves in the $0\xi\eta$-plane, using equations 1, 4, and 5, is to study the position of working point $M(a_1; a_0)$ in the same plane. For sampled-data control systems with transport lag, one great advantage of the described method is its ability to provide correlation between over-all system performance and time T_t.

Absolute Stability

Attainment of absolute stability is now sought. The first step is to extend the analysis procedure to learn how transport lag and other parameters will affect the stability of the system. A simple example will be presented but the same principles can be used with minor modifications, for other configurations.

The representative system shown in Fig. 1 has the transfer function

$$G(s)=\frac{Ka}{s(s+a)} \quad (6)$$

Applying equation 3, the transform $G(z, T_t)$ corresponding to this system is

$$G(z, T_t)=z^{-(k-1)}K\times$$
$$\left[\frac{1}{z-1}-\frac{1}{z-e^{-aT}}e^{-a(kT-T_t)}\right] \quad (7)$$

When the time T_t is less than one sampling period T or

$$k=1 \quad (8)$$

By substituting this value in equation 7 the roots of characteristic equation $1+G(z, T_t)=0$ are determined by

$$z^2+(K[1-e^{-a(T-T_t)}]-[1+e^{-aT}])z+$$
$$Ke^{-a(T-T_t)}+e^{-aT}(1-K)=0 \quad (9)$$

As known, the system is stable if all roots of the characteristic equation in z lie within the unit circle around the z-plane's origin.

The foregoing statement finds expression in terms of analysis in the $0\xi\eta$-plane if the limits of the stable region in that plane are determined by mapping the unit circle ($\omega_z=1$) from the z-plane into the $0\xi\eta$-plane.

The stable region in the $0\xi\eta$-plane of the representative system is bounded by

1. The locus of points (Γ^1) corresponding to the roots with constant $\omega_0=1$, which is the limit curve for complex roots.

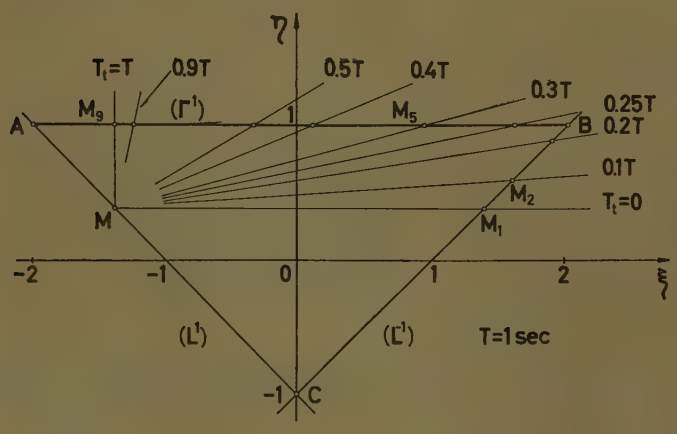


Fig. 3. Determination of gain K_{max} for $0\leq T_t\leq T=1$ sec

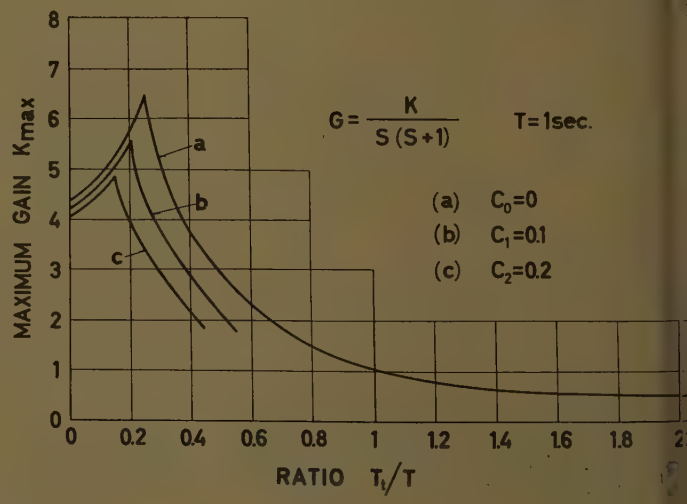


Fig. 4. Diagram relating gain K_{max} to ratio T_t/T for various values of constant c

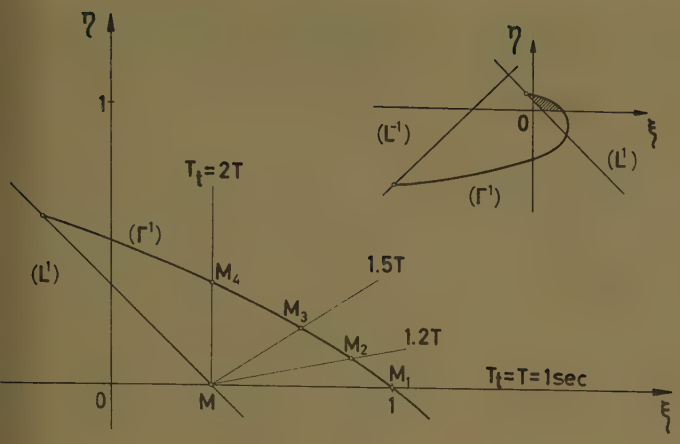


Fig. 5. Determination of gain K_{max} for $T\leq T_t\leq 2T$, $T=1$ sec

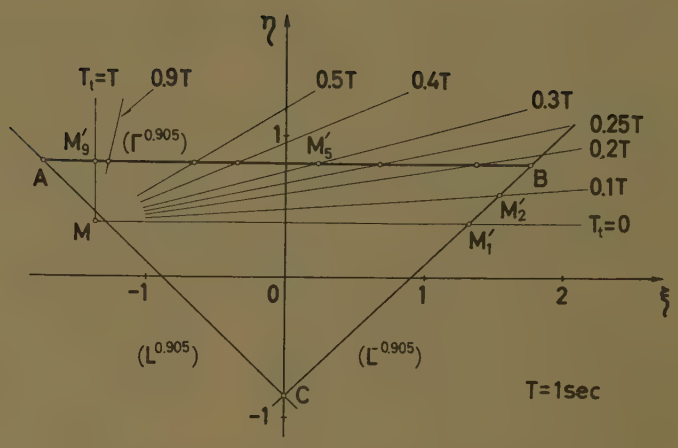


Fig. 6. Determination of gain K_{max} for $c_1=0.1$

Since

$-\omega_n\xi T=\omega_z=1$

nd

$$\zeta_z=-\cos\omega_n T\sqrt{1-\zeta^2} \tag{10}$$

he equations 5 become

$$\left.\begin{aligned}&=\phi_2(\zeta_z)\\ &=-\phi_1(\zeta_z)\end{aligned}\right\}\ (\Gamma^1) \tag{11}$$

Because

$\phi_1(\zeta_z)=-1$

$\phi_2(\zeta_z)=2\zeta_z$

see reference 3), equations 11 of the locus f points, corresponding to the roots with onstant $\omega_z=1$, obtain the form

$$\left.\begin{aligned}&=2\zeta_z\\ &=1\end{aligned}\right\}\ (\Gamma^1) \tag{12}$$

where ζ is the variable parameter. As varies only between -1 and $+1$, the quations 12 represent the straight line etween A and B in Fig. 2.

The locus of points (L^1), corresponding o positive unit roots $(z=+1)$, which is he limit curve for positive roots.

The locus (L^1) may be written simply by etting $z=+1$ in equation 1 and by taking oefficients a_1 and a_0 as variables ξ and η. o, in the case of equation 9

$$\eta=-\xi-1 \qquad (L^1) \tag{13}$$

hich represents line (L^1) in Fig. 2.

The locus of points (L^{-1}), correspond- ng to negative unit roots $(z=-1)$, which the limit curve for negative roots.

The locus (L^{-1}) is written in a form milar to that of (L^1)

$$\eta=\xi-1 \qquad (L^{-1}) \tag{14}$$

his represents line (L^{-1}) in Fig. 2.

Now, the condition of absolute tability, in terms of $0\xi\eta$-plane analysis, is

$$\eta<(K[1-e^{-a(T-T_t)}]-[1+e^{-aT}];\quad Ke^{-a(T-T_t)}+e^{-aT}(1-K))$$

nd is inside the stable region bounded by the curves (Γ^1), (L^1), and (L^{-1}) in Fig. 2.

Note that co-ordinates of working point M are functions of system parameters a, K, T, and T_t. Thus, by plotting in the same $0\xi\eta$-plane, the stable region and loci of working points M for different values of system parameters, the influence of each parameter variation upon system stability is easily determined.

When gain K and sampling period T are held constant, correlation between parameter a and transport lag T_t is found by first assuming that $K=2$ and $T=1$ sec (second).

Substituting these values of K and T in co-ordinates of working point M gives

$$\left.\begin{aligned}\xi&=1-2e^{-a(1-T_t)}-e^{-a}\\ \eta&=2e^{-a(1-T_t)}-e^{-a}\end{aligned}\right\} \tag{15}$$

For various values of a and T_t, equations 15 represent the net of two families of curves as shown in Fig. 2. From the part of this net within the triangular ABC or stable region, interpolation can be employed in finding any pair of a and T_t values which makes the system stable.

Attention also may be focused on correlating gain K and transport lag T_t. For example, suppose $a=1$ and $T=1$ sec, the relation between maximum gain for stability limit K_{max} and transport lag T_t is to be found. By inserting the given values for a and T into equations 15, the co-ordinates of working point M become

$$\left.\begin{aligned}\xi&=[1-e^{-(1-T_t)}]K-1.368\\ \eta&=[e^{-(1-T_t)}-0.368]K+0.368\end{aligned}\right\} \tag{16}$$

Evidently, if T_t is held constant and K is allowed to vary, the stable region remains unchanged, but working point M moves along the straight line in the $0\xi\eta$-plane determined by equations 16. For various values of T_t, equations 16 represent the family of straight lines plotted in Fig. 3. Now, from the intersection points (M_1, M_2, ... M_9) of these straight lines and the lines limiting the stable region, by using one of equations 16, the corresponding values of gain K_{max} are determined graphically.

In the simple example, analytical solution is also possible. It consists of determining the intersection points (M_1, M_2, ... M_9) with the help of equations 12, 14, and 16. When the co-ordinates ξ and η of these points are obtained, the corresponding equation in equations 16 will yield the gain values K_{max}. In doing so, the first part of the curve a in Fig. 4, which corresponds to the case when T_t is less than T, can be obtained.

So far, the analysis has been carried out for the case $0\leqslant T_t\leqslant T$. The second part of curve a, which corresponds to the second sampling period, is given by first inserting into equation 7

$$k=2 \tag{17}$$

Now, for $T=1$ sec and $a=1$, the corresponding characteristic equation is

$$z^3-1.368z^2+(K[1-e^{-(2-T_t)}]+0.368)z+K[e^{-(2-T_t)}-0.368]=0 \tag{18}$$

Using the same procedure as for system equation 9, the curves which define the stable region, corresponding to system equation 18, are determined by

$$\left.\begin{aligned}\xi&=-1.368\phi_2(\zeta_z)+\phi_3(\zeta_z)\\ \eta&=1.368\phi_1(\zeta_z)-\phi_2(\zeta_z)\end{aligned}\right\}\ (\Gamma^1)$$

$$\eta=-\xi+0.368 \qquad (L^1)$$

$$\eta=\xi+2.368 \qquad (L^{-1}) \tag{19}$$

The loci of points (Γ^1), (L^1), and (L^{-1}) have the form shown in the upper right corner in Fig. 5.

From system equation 18, the co-ordinates of working point M are

$$\left.\begin{aligned}\xi&=[1-e^{-(2-T_t)}]K+0.368\\ \eta&=[e^{-(2-T_t)}-0.368]K\end{aligned}\right\} \tag{20}$$

Both the enlarged part of the stable region and the straight lines defined by equations 20 for different values of T_t are shown in Fig. 5. Applying the same procedure used for the first sampling period, the second part of curve a is plotted in Fig. 4. The computation necessary to produce such diagrams for more complicated configurations can be performed by the proposed design procedure with relative ease and in a systematic manner. In designing sampled-data control systems, where transport lag is introduced to provide stabilization, the proposed procedure is extremely useful.

Where stabilization and compensation by conventional means are difficult, sampling can be used often—if the sampling period T is properly chosen—to improve stability. Thus, when applying the proposed technique, an optimum value of the sampling rate can be designed without creating new problems. Only sampling period T and gain K need to be considered as variable parameters. By investigating positions of working point M, as in the foregoing examples, suitable magnitude of the sampling rate can be determined easily.

Settling Time

In sampled-data control systems, the prescribed maximum settling time often appears as a requirement. This usually signifies that the adjustable parameters of a given system must be set so that the system will settle to an essentially steady-state value in the required time after application of a step-input function.

Settling time will not exceed the prescribed maximum value if all poles of the system-transfer function have a real part greater than some corresponding constant value c; that is, if

$$\omega_n\zeta \geq c$$

Because of equation 4, the foregoing condition means that all zeros of the characteristic equation in z should lie within the circle whose radius is

$$\omega_z = e^{-\omega_n\zeta T} = e^{-cT} \qquad (21)$$

Since this condition is so similar to that of absolute stability ($\omega_z = 1$), the design procedure which determines a system with a prescribed settling time is essentially that used in the preceding section. The only difference is that the radius of the circle in the z-plane is not unity.

To illustrate the design procedure, consider again the system shown in Fig. 1 with the transfer function $G(s)$ given in equation 6. Suppose parameter values are again $a=1$ and $T=1$ sec, and values are required for maximum gain K_{max} and transport lag T_t, which will result in a settling time corresponding to $c_1=0.1$.

Since the system will be investigated for T_t values that are less than the sampling period T—that is, for $k=1$—the corresponding characteristic equation is 9, into which is inserted $a=1$ and $T=1$ sec, giving the following equation

$$z^2+(K[1-e^{-(1-T_t)}]-1.368)z+ K(e^{-(1-T_t)}-0.368]+0.368=0 \qquad (22)$$

Settling time will be less than the prescribed maximum value determined by $c_1=0.1$ if the roots of equation 22 lie inside the circle whose radius is

$$\omega_z = e^{-c_1T} = 0.905 \qquad (23)$$

Limiting curves in the $0\xi\eta$-plane, corresponding to this circle, are obtained by the same method described in the foregoing section. Curves ($\Gamma^{0.905}$), ($L^{0.905}$), and ($L^{-0.905}$) consequently are defined by

$$\xi = a_2\phi_2(\zeta_z)e^{-c_1T} = 1.81\zeta_z$$

$$\eta = -a_2\phi_1(\zeta_z)e^{-2c_1T} = 0.82 \quad (\Gamma^{0.905})$$

$$\eta = e^{-c_1T}\xi - a_2e^{-2c_1T} = 0.905\xi - 0.82 \quad (L^{-0.905})$$

$$\eta = -e^{-c_1T}\xi - a_2e^{-2c_1T} = -0.905\xi - 0.82 \quad (L^{0.905}) \qquad (24)$$

These curves are shown in Fig. 6.

Now, in terms of $0\xi\eta$-plane analysis, the condition for maximum settling time may be formulated. Settling time will be less than the prescribed value corresponding to $c_1=0.1$, if the working point

$$M(K[1-e^{-(1-T_t)}]-1.368; \quad K[e^{-(1-T_t)}-0.368]+0.368)$$

is bounded by curves determined in equations 24.

Since the co-ordinates of point M are the same as in equations 16, the system of straight lines plotted in Fig. 6 is the same as was plotted in Fig. 3. The same graphical procedure used in the foregoing section will produce curve b in Fig. 4 from points M_1', M_2', ... M_9'. In like manner, if the required maximum settling time corresponds to $c_2=0.2$, curve c in Fig. 4 can be plotted.

Bearing in mind that curve a corresponds to $c_0=0$—that is, $\omega_z=1$—the tendency for gain K_{max} to change is realized, the amount depending on the magnitude of the maximum settling time prescribed and the transport lag. By using interpolation, diagrams like Fig. 4 are useful in determining the pairs of parameter values K and T which will yield the desired settling time. Since similar diagrams are obtainable when the relative damping is specified (considering in equations 5 relative damping coefficient ζ as constant),[3] the proposed method is applicable in designing systems of transient performance. For systems in which hidden oscillations[5,6] can take place, the results obtained in the design of transient behavior by this procedure should be taken with some caution.

Nonsynchronized Samplers

Another application of the proposed method is to systems containing two or more samplers operating with the same sampling rates, but not synchronized in phase, commonly known as sampled-data systems with nonsynchronized samplers.[4]

The block diagram in Fig. 7(A) represents the simplest type, having only two samplers S and S_Δ with the same sampling period T. However, S_Δ "slips" behind S by ΔT sec, where Δ is a fraction, and may be referred to as the slip factor. Based on the concept of equivalent samplers,[4] S_Δ may be represented by basic sampler S, preceded by an advance of ΔT sec and followed by a delay of ΔT sec, as in Fig. 7(B).

With reference to Fig. 7(B), the open-loop z-transfer function is

$$G(z,\Delta) = \mathfrak{Z}[e^{\Delta Ts}G_1(s)]\mathfrak{Z}[e^{-\Delta Ts}G_2(s)] \qquad (25)$$

where the notation $\mathfrak{Z}$ indicates the z-transform corresponding to the function which it prefixes. When $G_1(s)$ and $G_2(s)$ are rational functions in s, it follows from equation 3 and the definition of the modified z-transform that

$$\mathfrak{Z}[e^{\Delta Ts}G_1(s)] = zG_1(z,m)|_{m=\Delta} = zG_1(z,\Delta)$$

$$\mathfrak{Z}[e^{\Delta Ts}G_2(s)] = G_2(z,m)|_{m=1-\Delta} = G_2(z,1-\Delta) \qquad (26)$$

Where the roots of the corresponding characteristic equation are located w

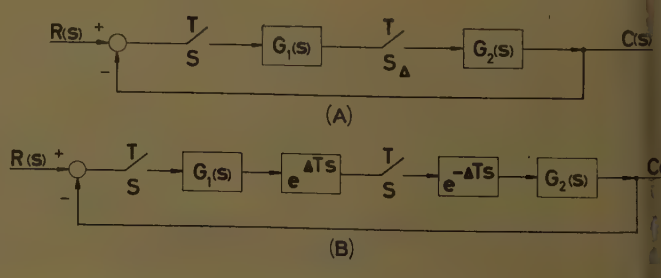


Fig. 7. A—Block diagram of basic nonsynchronized sampled-data system. B—Equivalent block diagram of system A

etermine stability of the system repre-
nted in Fig. 7(A); see equation 27,
hich shows the influence of the slip
ctor.

$$+zG_1(z, \Delta)G_2(z, 1-\Delta)=0 \quad (27)$$

When the values of Δ fall within a
rtain range, the nonsynchronized sam-
ling usually provides a stabilizing effect.
n a distinct class of sampled-data
ystems, nonsynchronized sampling is
troduced deliberately to improve stabil-
y. In designing such systems, the
ip factor's optimum value frequently
ust be determined to find the highest
tability limit compatible with maximum
llowable gain. To illustrate use of the
roposed technique, a simple numerical
xample is presented, based on the block
iagram, Fig. 7(A). Transfer functions
$G_1(s)$ and $G_2(s)$ are

$$G_1(s)=\frac{K(1-e^{-sT})}{s(s+1)} \quad G_2(s)=\frac{1}{s+0.5} \quad (28)$$

or both samplers, the sampling period
$T=1$ sec. The effect of slip factor
on system stability is now studied.

Making use of equations 25, 26, and
7, the corresponding characteristic equa-
ion is

$$z^2-(0.97-K[e^{-0.5(1-\Delta)}-e^{-0.5(1+\Delta)}])z+K[e^{-0.5(1+\Delta)}-0.368e^{-0.5(1-\Delta)}]+0.22=0 \quad (29)$$

Since the shape of characteristic curves
Γ^1), (L^1), and (L^{-1}) does not depend
n values of coefficients a_1 and a_0, equa-
ion 1, the stable region corresponding
o equation 29 is the same as that shown
n Figs. 2 and 3, which corresponds to
he characteristic equation 9.

Co-ordinates of working point M are

$$=[e^{-0.5(1-\Delta)}-e^{-0.5(1+\Delta)}]K-0.97$$
$$=[e^{-0.5(1+\Delta)}-0.368e^{-0.5(1-\Delta)}]K+0.22 \quad (30)$$

For various values of slip factor Δ
nd gain constant K, a system of straight
ines similar to that shown in Fig. 3 can
e constructed. Then, using the same
graphical procedure as was applied to Fig. 3, diagram (*a*) in Fig. 8 is plotted. Now, obviously, any value of slip factor Δ, which lies between 0 and 1, will improve system stability. The optimum value, procured from this diagram, is 0.76.

If transfer function $G_1(s)$ has the form

$$G_1(s)=\frac{Ke^{-sT}}{(s+1)^2+1} \quad (31)$$

and transfer function $G_2(s)$ remains unchanged, the corresponding characteristic equation is

$$z^3-0.753z^2+0.607Ke^{-1.5\Delta}\sin\Delta+0.106)z+0.082Ke^{-1.5\Delta}\sin(1-\Delta)-0.01=0 \quad (32)$$

In like manner, by applying the proposed technique, diagram (*b*) in Fig. 8 is evolved. In contrast to the preceding case, where nonsynchronized sampling provided a stabilizing effect, here the introduction of any slip factor value decreases the stability margin.

Plotting of diagrams that correspond to those in Fig. 8 reveals whether deterioration or improvement will result from nonsynchronized sampling in a given system, and shows the optimum value of slip factor Δ. The computation is relatively easy. Diagrams relating Δ to other parameters can also be plotted. They prove quite useful when designing sampled-data control systems where nonsynchronized sampling occurs inherently and where it is required for stabilizing effects.

Conclusions

The proposed design procedure is a graphical means of analyzing and synthesizing sampled-data control systems in which transport lag occurs inherently or is introduced deliberately to provide stabilization. It consists of plotting some characteristic curves in the $0\xi\eta$-plane and then investigating the positions of working point M as the function of system parameters. In most cases, all

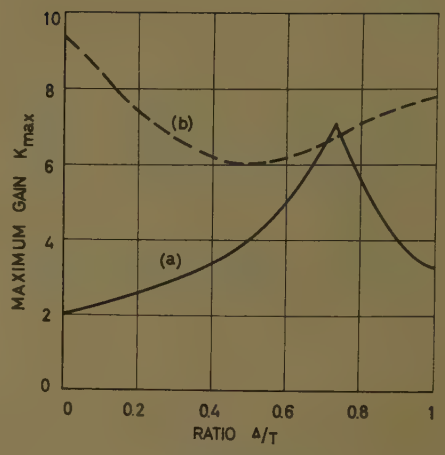


Fig. 8. Diagrams relating gain K_{max} to ratio Δ/T for two different systems

corresponding curves need not be pl tted, and often the determination of only part of a curve is necessary (as shown in Fig. 5). Computations can be performed with ease.

The technique is applicable to many varied problems that appear in sampled-data systems with two or more samplers. Analysis and synthesis of such systems, having finite sampling duration, can also be performed in a straightforward manner.

References

1. Graphical Analysis and Synthesis of Feedback Control Systems, I—Theory and Analysis, Dušan Mitrović. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 77, 1958 (Jan. 1959 section), pp. 476–87.

2. Graphical Analysis and Synthesis of Feedback Control Systems, II—Synthesis, Dušan Mitrović. *Ibid.*, pp. 487–96.

3. Graphical Analysis and Synthesis of Feedback Control Systems, III—Sampled-Data Feedback Control Systems, Dušan Mitrović. *Ibid.*, pp. 497–503.

4. Digital and Sampled-Data Control Systems (book), J. T. Tou. McGraw-Hill Book Company, Inc., New York, N. Y., 1959.

5. Sampled-Data Control Systems (book), Eliahu I. Jury. John Wiley & Sons, Inc., New York, N. Y., 1958.

6. Hidden Oscillations in Sampled-Data Control Systems, Eliahu I. Jury. *Ibid.*, vol. 75, 1956 (Jan. 1957 section), pp. 391–95.

7. The Analysis of Sampled-Data Systems, J. R. Ragazzini, L. A. Zadeh. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 71, Nov. 1952, pp. 225–34.

Discussion

M. Anantha Pai (University of California, Berkeley, Calif.): This paper is a useful contribution to the literature in that it extends Mitrović's graphical analysis[1] to linear sampled-data systems with time lag. Using the criteria for stability of linear discrete systems recently discussed by Jury[2–4] provides an alternate analytical means of deriving the same results as reported in the paper under discussion. In general terms, the new criteria achieve the same degree of simplicity for sampled-data systems as does the Lienard-Chipart criteria[5] (which are a further simplification of Routh–Hurwitz criteria) for continuous systems. These will be stated now for the second- and third-order case and then used to find boundaries of the stable region in the $0\eta\xi$-plane as well as values for K_{max}. Criteria for the general nth-order case are contained in reference 4.

Stability Test, Second Order Case[4]

where $n=2$, $F(z)=a_2z^2+a_1z+a_0$, $a_2>0$

$$|a_0|<a_2$$
$$|a_0+a_2|>|a_1| \text{ or } a_0+a_1+a_2>0$$
$$a_0-a_1+a_2>0 \quad (33)$$

Stability Test, Third Order Case[4]

where $n=3$, $F(z)=a_3z^3+a_2z^2+a_1z+a_0$, $a_3>0$

$$|a_0|<a_3$$
$$a_0^2-a_3^2<a_0a_2-a_1a_3$$
$$a_0+a_1+a_2+a_3>0$$
$$a_0-a_1+a_2-a_3<0 \quad (34)$$

Example 1

Considering characteristic equation 9 of the paper, recognizing that $a_0=\eta$, $a_1=\xi$ and using equations 33 readily produces

$$|\eta|<1 \quad (35)$$

$$\eta+1+\xi>0 \quad (36)$$

$$\eta-\xi+1>0 \quad (37)$$

The three inequalities give the stability region of Fig. 3 of this paper.

EXAMPLE 2

Considering equation 18 of the paper, which is a polynomial in z of third degree, and using equation 34 will give the following:

$$|\eta|<1 \quad (38)$$

$$\eta^2-1<-1.368\,\eta-\xi \quad (39)$$

$$\eta+\xi-1.368+1>0 \quad (40)$$

$$\eta-\xi-1.368-1<0 \quad (41)$$

Careful interpretation of these inequalities will reveal the stability region shown in Fig. 5 of the paper.

EXAMPLE 3

Consider an example of synthesis to find the K_{max} required for a given system to remain stable. The stability test of equation 32 of the paper is

$$|0.082Ke^{-1.5\Delta}\sin(1-\Delta)-0.01|<1 \quad (42)$$

$$(0.082Ke^{-1.5\Delta}\sin(1-\Delta)-0.01)^2-1<-0.753(0.082Ke^{-1.5\Delta}\sin(1-\Delta)-0.01)-(0.607Ke^{-1.5\Delta}\sin\Delta+0.106) \quad (43)$$

$$0.082Ke^{-1.5\Delta}\sin(1-\Delta)-0.01+0.607Ke^{-1.5\Delta}\sin\Delta+0.106-0.753+1>0 \quad (44)$$

$$0.082Ke^{-1.5\Delta}\sin(1-\Delta)-0.01-0.607Ke^{-1.5\Delta}\sin\Delta-0.106-0.753-1<0 \quad (45)$$

For a given value of Δ, relations 42 through 45 determine the allowable values of K. Thus, K_{max} is readily obtained from the inequalities. For example, take $\Delta=1$ corresponding to $\Delta/T=1$ in Fig. 8, where equations 42–45 yield relations 46–49,

$$0.01<1 \quad (46)$$

This gives no information on K.

$$K<7.93 \quad (47)$$

$$K>-3.07 \quad (48)$$

$$K>-14.7 \quad (49)$$

Since only positive values of K are considered, $K_{max}=7.93$, which agrees with the value corresponding to $\Delta/T=1$ in curve (b) of Fig. 8. Similarily, K_{max} could also be obtained for other values of Δ. This method could be applied also to examples contained in reference 5 of the paper. Stability tests in reference 4 will give the criteria for absolute stability. In the case of instability, it also indicates the number of roots which lie inside or outside the unit circle.

While the paper gives a useful graphical design technique, the present discussion outlines an alternate method which achieves the same results by using Jury's simplified stability criteria.[4] Thus, the design engineer has a choice.

REFERENCES

1. See references 1, 2, and 3 of the paper.

2. NOTES ON THE STABILITY CRITERION FOR LINEAR DISCRETE SYSTEMS, Eliahu I. Jury, B. H. Bharucha. *Transactions*, Professional Group on Automatic Control, Institute of Radio Engineers, New York, N. Y., vol. AC-6, no. 1, Feb. 1961, pp. 89–90.

3. ADDITIONS TO STABILITY CRITERION FOR LINEAR DISCRETE SYSTEMS, Eliahu I. Jury. *Ibid.*, vol. AC-6, no. 1, Sept. 1961, pp. 342–43.

4. A SIMPLIFIED STABILITY CRITERION FOR LINEAR DISCRETE SYSTEMS, Eliahu I. Jury. *Report no. 373*, Electronics Research Laboratory, University of California, Berkeley, Calif., Series no. 60, June 1961.

5. *Theory of Matrices* (book), F. R. Gant Macher, Chelsea Publishing Co., New York, N. Y., vol. I, 1959, pp. 221–27.

Dragoslav Šiljak: I read with interest the discussion by M. Anantha Pai and do not consider additional comment necessary.

Mathematical Models for Time-Domain Design of Electrohydraulic Servomechanisms

P. K. C. WANG
ASSOCIATE MEMBER AIEE

IN THE ANALYTICAL DESIGN of high-speed electrohydraulic servomechanisms, an accurate mathematical model of the valve-actuator mechanism is generally helpful in determining the controller and system configuration required to achieve the desired over-all performance. During past years, various forms of electrohydraulic valves and actuator mechanisms have been realized. Detailed descriptions of these devices can be found in a book by Blackburn et al.[1] Most previous works on the analysis of these systems have resorted to linearization techniques.[2,3] Recently, Butler and Turnbull presented detailed analysis of simplified nonlinear models.[4,5] Also, Zaborszky and Harrington derived describing functions for both single and two-stage electrohydraulic valves, which permit analytical design in the frequency-domain.[6–9] The limitations of linearized models are discussed by Rausch.[10]

The objective of this paper is to derive mathematical models which will represent most conventional electro-hydraulic valve-controlled actuators with sufficient accuracy for time-domain design and to examine the system from a general viewpoint so as to reveal some of its salient features. Emphasis is placed upon derivation of time-domain trajectories for various types of inputs. The results form a basis for further investigations on sampled-data and time-optimal control of electrohydraulic systems.

Paper 61-819, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Ithaca, N. Y., June 18–23, 1961. Manuscript submitted October 14, 1960; made available for printing April 21, 1961.

P. K. C. WANG is with the International Business Machines Corporation, San Jose, Calif.

The author wishes to thank J. T. Ma and J. O. Hildebrand for their many helpful discussions.

System Description

The system under consideration is shown in Fig. 1. The functional parts consist of a 4-way valve with a generalized valve-spool driver and a piston-type actuator rigidly attached to the load. The valve-spool driver is usually an electric torque-motor which is either connected directly to the spool or indirectly through additional stages of hydraulic

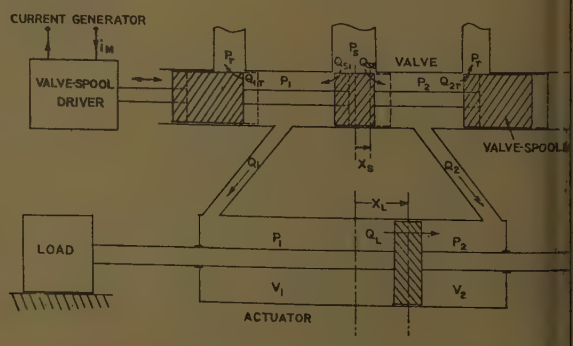


Fig. 1. A typical electro-hydraulic valve-controlled actuator

plification. The electric torque-motor ssumed to be driven by an ideal cur- t generator so as to eliminate the effect equivalent motor armature series im- lance. Also, the actuator load con- of mass M_L and a general friction load

The valve-spool displacement $x_s(\tau)$ is asured from a neutral position where spool is symmetrically centered with pect to the valve ports. The actuator ton displacement, $x_L(\tau)$, is also defined m a fixed reference corresponding to the ddle point of the actuator cylinder. th $x_s(\tau)$ and $x_L(\tau)$ are chosen to be sitive from left to right.

sic Assumptions

The valve is symmetrical (that is, the et and outlet ports are identical).

The oil in the valve cylinder is in- npressible and the flow is steady.

The volumetric flow Q through the ve ports can be related to the pressure ferential ΔP and port area A_p instan- neously by

$$Q = C_v A_p \operatorname{sgn} \Delta P \sqrt{|\Delta P|} \qquad (1)$$

ere C_v is an average discharge propor- nality constant.

The valve is located near the actuator that pressure drops and propagation ne lag due to connecting lines are neg- ible. Also, the pressures are uniform ide the actuator cylinder volumes V_1 d V_2.

The supply pressure is maintained at constant value P_s.

Effects of temperature variations and ing on the oil properties are negligible.

Further assumptions will be mentioned, ere applicable, in the ensuing discussion.

odel for Large Motions

The equation of motion for the valve ool is

$$M_s\ddot{x}_s(\tau)+F_s[\dot{x}_s(\tau)]+K_s x_s(\tau) = F_R+F_T[i_M(\tau)] \qquad (2)$$

where M_s is the spool mass, F_s is a friction force function depending upon the spool velocity, K_s is a valve spring constant, and F_T is assumed to be an amplitude-sensitive nonlinear function of the actual torque-motor current $i_M(\tau)$. In particular, F_T may be a hysteresis nonlinearity caused by magnetic hysteresis in the motor. F_R is the axial component of Bernoulli reaction force which is induced by a change in the momentum of oil discharging through the valve ports. Lee and Blackburn have shown that F_R can be decomposed into steady-state and transient components.[11] Both components are functions of the flow rate and valve port geometry. Thus, coupling exists between the motions of the spool and actuator load. However, in practical design, the actuating force on the valve spool can be made large compared with F_R by adequate choice of torque-motor. Furthermore, proper valve design permits reduction of Coulomb and stiction forces to some nominal values. Since the spool motion is small, equation 2 can be approximated by a second-order linear differential equation of the form:

$$\ddot{x}_s(\tau)+2\zeta_v\omega_{nv}\dot{x}_s(\tau)+\omega_{nv}^2 x_s(\tau) = \frac{F_T[i_M(\tau)]}{M_s} \qquad (3)$$

The damping ratio ζ_v and natural frequency ω_{nv} may be estimated from the manufacturers' valve frequency response data.

The flow through the valve ports is derived by applying equation 1 and the continuity relationships; see Fig. 1:

$$Q_1 = Q_{s1} - Q_{1r} \qquad (4)$$

$$Q_2 = Q_{s2} - Q_{2r} \qquad (5)$$

By considering the mass flow rate into the actuator cylinder volumes V_1 and V_2, relationships between Q_1, Q_2, and the load motion are established.

$$Q_1 = \frac{V_1(\tau)}{\beta}\dot{P}_1(\tau)+\dot{V}_1(\tau)+Q_L \qquad (6)$$

$$Q_2 = \frac{V_2(\tau)}{\beta}\dot{P}_2(\tau)+\dot{V}_2(\tau)-Q_L \qquad (7)$$

where β is the bulk modulus of oil. Since the actuator volumes $V_1(\tau)$ and $V_2(\tau)$ vary linearly with load displacement—that is $V_1(\tau)=[V_T/2+A_a x_L(\tau)], V_2(\tau)=[V_T/2-A_a x_L(\tau)]$—equations 6 and 7 reduce to:

$$Q_1 = \frac{\left[\frac{V_T}{2}+A_a x_L(\tau)\right]}{\beta}\dot{P}_1(\tau)+\dot{x}_L(\tau)A_a+Q_L \qquad (8)$$

$$Q_2 = \frac{\left[\frac{V_T}{2}-A_a x_L(\tau)\right]}{\beta}\dot{P}_2(\tau)-\dot{x}_L(\tau)A_a-Q_L \qquad (9)$$

where V_T is the total oil volume inside the actuator cylinder.

Also, the leakage flow rate Q_L across the actuator piston can be approximated by a steady flow in an annulus between a finite length circular shaft and a concentric cylinder. In this case, Q_L is proportional to the pressure differential across the piston.

$$Q_L = C_L[P_1(\tau)-P_2(\tau)] \qquad (10)$$

Finally, the equation for load motion is

$$M_L\ddot{x}_L(\tau) = A_a[P_1(\tau)-P_2(\tau)]- F_L[\dot{x}_L(\tau)]+F_D(\tau) \qquad (11)$$

$F_D(\tau)$ is an external load disturbance force.

The complete system equations can be conveniently written in the following vector form:

$$\frac{d\mathbf{U}}{d\tau} = \mathbf{H}(\mathbf{U})+\mathbf{\Psi}(\tau) \qquad (12)$$

where $\mathbf{U}(\tau)$ is the system state vector given by

$$\mathbf{U}(\tau) = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} = \begin{bmatrix} x_s(\tau) \\ \dot{x}_s(\tau) \\ x_L(\tau) \\ \dot{x}_L(\tau) \\ P_1(\tau) \\ P_2(\tau) \end{bmatrix} \qquad (13)$$

$\mathbf{\Psi}(\tau)$ is a vector forcing function

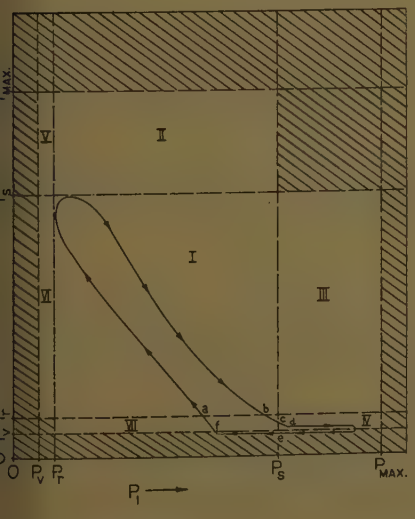


Fig. 2 (left). Discontinuous boundaries in pressure (P_1, P_2) plane

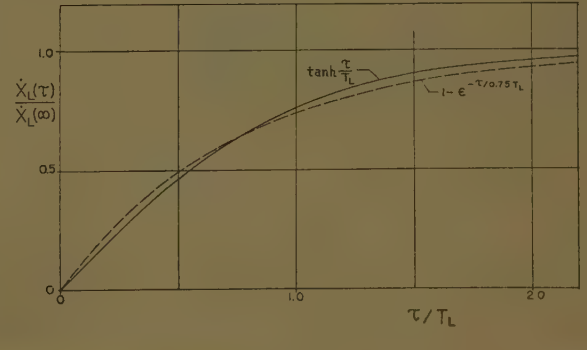


Fig. 3 (right). Normalized velocity response for a relaxed, frictionless system with step-current input

$$\Psi(\tau)=\begin{bmatrix}0\\ F_T[i_M(\tau)]/M_s\\ 0\\ F_D(\tau)/M_L\\ 0\\ 0\end{bmatrix} \quad (14)$$

and $\mathbf{H}(\mathbf{U})$ is a vector whose components h_1 are

$$h_1=\dot{x}_s(\tau) \quad (15)$$

$$h_2=-2\zeta_v\omega_{nv}\dot{x}_s(\tau)-\omega_{nv}^2x_s(\tau) \quad (16)$$

$$h_3=\dot{x}_L(\tau) \quad (17)$$

$$h_4=\frac{A_a}{M_L}[P_1(\tau)-P_2(\tau)]-\frac{F_L[\dot{x}_L(\tau)]}{M_L} \quad (18)$$

For "zero-lapped" valve

$$x_s(\tau)>0\begin{cases}h_5^+=\dfrac{\beta}{\left[\dfrac{V_T}{2}+A_ax_L(\tau)\right]}\{C_v|A_p[x_s(\tau)]|\ \text{sgn}\ [P_s-P_1(\tau)]\sqrt{|P_s-P_1(\tau)|}-\\ \qquad A_a\dot{x}_L(\tau)+C_L[P_1(\tau)-P_2(\tau)]\} \quad (19)\\ h_6^+=\dfrac{\beta}{\left[\dfrac{V_T}{2}-A_ax_L(\tau)\right]}\{-C_v|A_p[x_s(\tau)]|\ \text{sgn}\ [P_2(\gamma)-P_r]\sqrt{|P_2(\tau)-P_r|}+\\ \qquad A_a\dot{x}_L(\tau)-C_L[P_1(\tau)-P_2(\tau)]\} \quad (20)\end{cases}$$

$$x_s(\tau)<0\begin{cases}h_5^-=\dfrac{\beta}{\left[\dfrac{V_T}{2}+A_ax_L(\tau)\right]}\{-C_v|A_p[x_s(\tau)]|\ \text{sgn}\ [P_1(\tau)-P_r]\sqrt{|P_1(\tau)-P_r|}-\\ \qquad A_a\dot{x}_L(\tau)+C_L[P_1(\tau)-P_2(\tau)]\} \quad (21)\\ h_6^-=\dfrac{\beta}{\left[\dfrac{V_T}{2}-A_ax_L(\tau)\right]}\{C_v|A_p[x_s(\tau)]|\ \text{sgn}\ [P_s-P_2(\tau)]\sqrt{|P_s-P_2(\tau)|}+\\ \qquad A_a\dot{x}_L(\tau)-C_L[P_1(\tau)-P_2(\tau)]\} \quad (22)\end{cases}$$

Hereafter, $(\ )^+$ and $(\ )^-$ notations will be used to indicate the polarity of $x_s(\tau)$. Equations for h_5 and h_6 for underlapped and overlapped valves are given in Appendix I.

The system trajectories are restricted to a partially closed region in phase space ($\mathbf{U}$ space) as a result of the following constraints induced by physical limitations and design specifications:

$$\left.\begin{array}{ll}|x_s(\tau)|\le x_{s(\max)}: & \text{The spool travel is limited by the valve's physical size.}\\ |\dot{x}_s(\tau)|\le\dot{x}_{s(\max)}: & \text{The spool velocity is limited by friction and motor saturation.}\\ |x_L(\tau)|\le x_{L(\max)}: & \text{The actuator piston travel is confined to the working length of the cylinder such that } x_{L(\max)}<V_T/2A_a.\\ |\dot{x}_L(\tau)|\le\dot{x}_{L(\max)}: & \text{For an external disturbance-free system, the peak actuator piston velocity is limited primarily by friction and supply pressure } P_s.\\ P_{\max}>P_1\ge P_v,\ P_{\max}>P_2\ge P_v: & \text{The instantaneous pressures should be kept below a specified safe pressure } P_{\max}. \text{ Assuming that the oil in the actuator cylinder is ideal and does not completely vaporize, } P_1 \text{ and } P_2 \text{ cannot be lower than the oil vapor pressure } P_v.\end{array}\right\} \quad (23)$$

By inspection of equation 12, it is evident that its right-hand side is discontinuous in nature. Moreover, the phase co-ordinates are bounded due to constraints given by equation 23. The discontinuities arise from the bidirectional flow property of the valve ports and nonsymmetry of the valve load.

The system equations may be rewritten in an alternative general vector form:

$$\frac{d\mathbf{U}}{d\tau}=\mathbf{H}_i(\mathbf{U})+\Psi(\tau),\quad i=1,2.....N \quad (24)$$

plus a set of mode boundaries

$$Z_k(\mathbf{U})=0,\quad k=1,2.....M \quad (25)$$

Equation 25 represents a set of hypersurfaces which partition the phase space into distinct regions Ω_j. Within each Ω_j, the system behavior is described by a particular equation in 24, whose $\mathbf{H}_i(\mathbf{U})$ is continuous in $\mathbf{U}$ everywhere within Ω_j.

For the system under consideration, the mode boundaries are

$$\begin{array}{l}x_s(\tau)=0 \text{ (for "zero-lapped" valve)}\\ P_s-P_1(\tau)=0\\ P_s-P_2(\tau)=0\\ P_1(\tau)-P_r=0\\ P_2(\tau)-P_r=0\\ \dot{x}_E(\tau)=0 \text{ (if Coulomb friction load is present)}\end{array} \quad (26)$$

Additional mode boundaries are introduced by the last two constraints in equation 23,

$$\begin{array}{l}P_1(\tau)-P_v=0\\ P_2(\tau)-P_v=0\end{array} \quad (27)$$

These represent the cavitation conditions. A detailed discussion of the cavitation phenomenon will be given elsewhere.

There are some basic questions pertaining to the behavior of the trajectories upon traversing the mode boundaries. For example, the pressures may change rapidly across certain boundaries. Such cases will not be discussed in the present paper.

From the preceding discussions, it is evident that the electrohydraulic actuator is basically a multiple-mode nonlinear system. This point is further clarified by Fig. 2, showing the projection of a typical phase-space trajectory onto the pressure (P_1, P_2) plane. The scales are grea exaggerated so as to reveal various mo regions. The permissible trajectories a confined within and to the boundar of the unshaded regions. Consider typical pressure phase trajectory, ge erated by certain forcing $\Psi(\tau)$, starti at a. Its mode sequence is as follows

Trajectory	Mode Region	Pressure Boun
ab	I	$P_r<P_1(\tau)\le P_s$, $P_r<P_2(\tau)\le$
bc	VII	$P_r<P_1(\tau)\le P_s$, $P_v<P_2(\tau)\le$
cd	IV	$P_s\le P_1(\tau)<P_m$, $P_v\le P_2(\tau)<$
de	Boundary of IV (cavitation)	$P_s\le P_1(\tau)<P_m$, $P_2(\tau)=P_v$
ef	Boundary of VII (cavitation)	$P_r<P_1(\tau)\le P_s$, $P_2(\tau)=P_v$
fa	VII	$P_r<P_1(\tau)<P_s$, $P_v\le P_2(\tau)\le$

Time-Domain Trajectories

In this section, analytical expressio for the time-domain trajectories of a c turbance-free system with "zero-lappe valve and negligible actuator leakage v be derived. The return pressure P_r is to zero since $P_s>>P_r$.

Response to Large-Step Motor-Current Inputs

The large-step motor-current respo is of particular interest to design of tim optimal electrohydraulic servomech nisms, where the maximum allowable m tor current is applied during the ini mode of operation.

Let the step input current be

$$i_M(\tau)=\begin{cases}0 \text{ for } \tau<0\\ i_o \text{ for } \tau>0\end{cases}$$

Since F_T is amplitude-sensitive, effective motor forcing is also a step w magnitude $F_T(i_o)$.

The steady-state equations for i_o are

$$x_s(\infty)=F_T(i_o)/\omega_{nv}^2M_s \quad (\ )$$

$$A_a\dot{x}_L(\infty)=C_vA_p[x_s(\infty)]\sqrt{P_s-P_1(\infty)} \quad (\ )$$

$$A_a\dot{x}_L(\infty)=C_vA_p[x_s(\infty)]\sqrt{P_2(\infty)} \quad (\ )$$

$$A_a[P_1(\infty)-P_2(\infty)]=F_L[\dot{x}_L(\infty)] \quad (\ )$$

For viscous friction load, $F_L[\dot{x}_L(\infty)]=f_L\dot{x}_L(\infty)$, the steady-state solutions h the form:

$$\dot{x}_L(\infty)=\frac{C_v^2A_p^2[x_s(\infty)]}{4A_a^2}\times\left\{\sqrt{\frac{f_L^2}{A_a^2}+\frac{8A_a^2P_s}{C_v^2A_p^2[x_s(\infty)]}}-\frac{f_L}{A_a}\right\} \quad (\ )$$

$$P_2(\infty)=\frac{A_a^2\dot{x}_L^2(\infty)}{C_v^2A_p^2[x_s(\infty)]} \quad (\ )$$

$$P_1(\infty)=P_s-P_2(\infty), \quad (\ )$$

where $x_s(\infty)$ is given by equation 28.

In the presence of both viscous and ulomb friction forces, that is, $F_L = {}_L + F_c \operatorname{sgn} \dot{x}_L$, the steady-state solu- ns are similar to equations 32, 33, and except for a replacement of P_s by $(P_s - /A_a)$.

Upon initiation of a step current, P_1 d P_2 change rapidly as a result of open- of the valve ports. Since the load not move instantaneously, the pres- e change is primarily due to compressi- ity of oil inside the actuator cylinder. tting $A_a\dot{x}_L = 0$ and $A_a x_L(\tau) = A_a x_L(0)$ equations 19 and 20 (for $i_o > 0$) gives proximate differential equations de- ibing the initial pressure buildup.

$$\dot{P}_1 \approx \frac{\beta}{\left[\frac{V_T}{2} + A_a x_L(0)\right]} \times \{C_v A_p[x_s(\tau)]\sqrt{P_s - P_1(\tau)}\} \quad (35)$$

$$\dot{P}_2 \approx \frac{\beta}{\left[\frac{V_T}{2} - A_a x_L(0)\right]} \times \{-C_v A_p[x_s(\tau)]\sqrt{P_2(\tau)}\} \quad (36)$$

The solutions are

$$(\tau) = P_s - \left\{(P_s/2)^{1/2} + \frac{\beta C_v}{[V_T + 2A_a x_L(0)]} \times \int_0^\tau A_p[x_s(t)]dt\right\}^2 \quad (37)$$

$$(\tau) = \left\{(P_s/2)^{1/2} + \frac{\beta C_v}{[V_T - 2A_a x_L(0)]} \times \int_0^\tau A_p[x_s(t)]dt\right\}^2 \quad (38)$$

The step current response of spool can approximated by the solution of uation 3.

$$(\tau) = \frac{F_T(i_0)}{\omega_{nv}^2 M_s}[1 - \epsilon^{-\zeta_v \omega_{nv}\tau} \times \cos \omega_{nv}\sqrt{1-\zeta^2}\tau] \quad (39)$$

Since the port area as a function of $x_s(\tau)$ known, $P_1(\tau)$, $P_2(\tau)$, $x_L(\tau)$, and $\dot{x}_L(\tau)$ ay be obtained by direct integration. nce the force acting on the load is A_a $P_1(\tau) - P_2(\tau)]$, equations 37 and 38 are eful for estimating the initial dead- ne in the step response caused by static ction.

After the pressures have reached their ak values, the oil flows are primarily ue to actuator piston motion, and the flow contributed by oil compressibility is small.

$$\frac{\left[\frac{V_T}{2} + A_a x_L(\tau)\right]}{\beta}\frac{dP_1}{d\tau} \approx 0,$$

$$\frac{\left[\frac{V_T}{2} - A_a x_L(\tau)\right]}{\beta}\frac{dP_2}{d\tau} \approx 0 \quad (40)$$

The system equations reduce to a Riccati equation:

$$\frac{d\dot{x}_L}{d\tau} = \frac{A_a}{M_L}\left\{P_s - 2\left[\frac{A_a\dot{x}_L(\tau)}{C_v A_p[x_s(\tau)]}\right]^2\right\} - \frac{f_L\dot{x}_L(\tau)}{M_L} \quad (41)$$

For a fast-acting valve with constant port width, W_o, $A_p[x_s(\tau)]$ quickly attains its steady-state value $W_o F_T(i_o)/\omega_{nv}^2 M_s$.

The subsequent system trajectories can be accurately described by the solution of equation 41:

$$\dot{x}_L(t+\tau) = \frac{C_v^2 W_o^2 x_s^2(\infty) M_L}{2A_a^3 T_L}\tanh\left[\frac{\tau}{T_L} + \tanh^{-1}\left(\frac{2A_a^3 T_L \dot{x}_L(t)}{C_v^2 W_o^2 x_s^2(\infty) M_L} + \frac{f_L T_L}{2M_L}\right)\right] - \frac{f_L C_v^2 W^2{}_o x_s^2(\infty)}{4A_a^3} \quad (42)$$

$$x_L(t+\tau) = x_L(t) + \frac{C_v^2 W_o^2 x_s^2(\infty) M_L}{2A_a^3}\left[\ln\left\{\frac{\cosh\left[\frac{\tau}{T_L} + \tanh^{-1}\left(\frac{2A_a^3 T_L \dot{x}_L(t)}{C_v^2 W_o^2 x_s^2(\infty) M_L} + \frac{f_L T_L}{2M_L}\right)\right]}{\cosh\left[\tanh^{-1}\left(\frac{2A_a^3 T_L \dot{x}_L(t)}{C_v^2 W_o^2 x_s^2(\infty) M_L} + \frac{f_L T_L}{2M_L}\right)\right]}\right\} - \frac{f_L}{2M_L}\tau\right] \quad (43)$$

$$P_1(t+\tau) = P_s - \frac{A_a^2}{C_v^2 W_o^2 x_s^2(\infty)}\dot{x}_L^2(t+\tau) \quad (44)$$

$$P_2(t+\tau) = P_s - P_1(t+\tau) \quad (45)$$

where

$$T_L = \frac{C_v W_o x_s(\infty) M_L}{2A_a^2} \times \left[\frac{P_s}{2} + \frac{f_L^2 C_v^2 W_o^2 x_s^2(\infty)}{16 A_a^4}\right]^{-1/2} \quad (46)$$

A normalized plot of load velocity given by equation 42 with $\dot{x}_L(t)$ and f_L equal to zero is shown in Fig. 3. For this case, the step response of a relaxed, frictionless system for a particular current i_o can be approximated by that of a one-integral plus one-time-constant system. The equivalent time constant T'_L is proportional to the input current amplitude i_o and has a value of approximately $0.75\ T_L$. Equations 42 and 43 are useful for quick estimation of the response time of a time-optimal electrohydraulic servomechanism for various step position inputs, since, in most cases, the load deceleration time is short as compared with the total response time.

Trajectories in Power Series Form

Power series extrapolations of system trajectories in the time domain have been used in the design of predictor control systems with linear dynamic processes.[12] This approach is especially useful in determining the required forcing functions if the control system inputs can be also extrapolated by power series in time. For nonlinear processes such as the electrohydraulic actuator considered in this paper, the use of a linearized model for accurate prediction of the time-domain trajectories is usually unsatisfactory. The power series approach is applicable.

Fig. 4 shows a block diagram of a proposed predictor electrohydraulic servomechanism. The system positional input $r(\tau)$ and the state variables u_t of the hydraulic actuator are measured periodically and fed into the computer as initial conditions for a particular computation period. The computer performs extrapolations of both the input, $r(\tau)$, and the actuator trajectories corresponding to certain assumed forcings for a future time τ. Then, on the basis of the extrapolated information and a predetermined forcing criterion, a control signal is generated and fed into the current driver.

In the subsequent discussion, power series solutions of equation 12 will be considered. Since the valve-spool motion is assumed to be independent of the remaining system dynamics, equation 12 can be written as a time-dependent equation by specifying the port area function as a power series in τ.

$$\frac{d\mathbf{Y}}{d\tau} = \mathbf{G}(\mathbf{Y}, \tau) \quad (47)$$

where

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} x_L(\tau) \\ \dot{x}_L(\tau) \\ P_1(\tau) \\ P_2(\tau) \end{bmatrix}$$

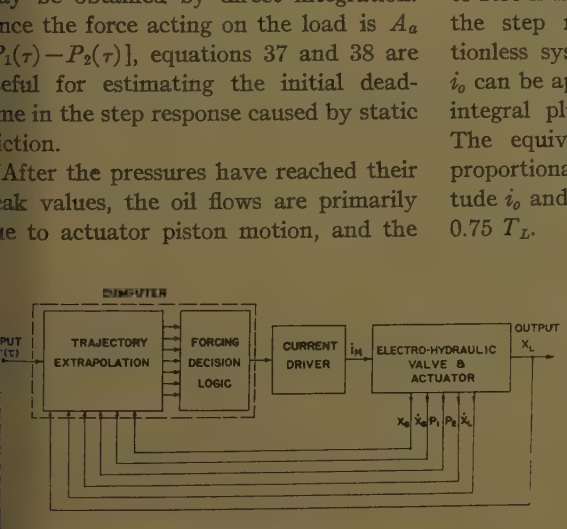


Fig. 4 (left). Block diagram of a proposed predictor electrohydraulic servomechanism

$$\mathbf{G}(\mathbf{Y},\tau)=\begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \end{bmatrix}=\begin{bmatrix} h_3 \\ h_4 \\ h_5^{\pm} \\ h_6^{\pm} \end{bmatrix} \quad (48)$$

$$\text{with } A_p[x_s(\tau)]=A_p(t)+\sum_{n=1}^{\infty}\alpha_n\tau^n \quad (49)$$

To determine power series solutions for equation 47, the system's operating mode region, Ω_j, in phase space, must be determined from the initial state $\mathbf{Y}(t)$ so that the appropriate form of $\mathbf{G}(\mathbf{Y},\tau)$ is used for computation. Within Ω_j, the components of $\mathbf{G}$(that is, $g_i(\mathbf{Y},\tau)$ are analytic functions of x_L, $\dot{x}_L$, P_1 and P_2 simultaneously. The general forms of power series solutions are

$$\left.\begin{aligned} x_L(t+\tau)&=x_L(t)+\sum_{n=1}^{\infty}\gamma_{1n}\tau^n \\ \dot{x}_L(t+\tau)&=\dot{x}_L(t)+\sum_{n=1}^{\infty}\gamma_{2n}\tau^n \\ P_1(t+\tau)&=P_1(t)+\sum_{n=1}^{\infty}\gamma_{3n}\tau^n \\ P_2(t+\tau)&=P_2(t)+\sum_{n=1}^{\infty}\gamma_{4n}\tau^n \end{aligned}\right\} \quad (50)$$

where

$$\gamma_{i1}=g_i(t)$$

$$\gamma_{i2}=1/2\left[\sum_{j=1}^{4}\frac{\partial g_i}{\partial y_j}\gamma_{j1}+\frac{\partial g_i}{\partial \tau}\right]\mathbf{Y}(\tau=0),\ \tau=0$$

$$\gamma_{i3}=1/3\left[\sum_{j=1}^{4}\frac{\partial g_i}{\partial y_j}\gamma_{j2}+1/2\sum_{j=1}^{4}\times\sum_{k=1}^{4}\frac{\partial^2 g_i}{\partial y_j\partial y_k}\gamma_{j1}\gamma_{k1}+\sum_{j=1}^{4}\times\frac{\partial^2 g_i}{\partial y_j\partial\tau}\gamma_{j1}+1/2\frac{\partial^2 g_i}{\partial\tau^2}\right]\mathbf{Y}_{(\tau=0),\,\tau=0}$$

$$\vdots \qquad\qquad \vdots$$

$$\gamma_{iN}=\frac{1}{N}P_{Ni}[\gamma_{j1},\ \gamma_{j2},\ \gamma_{j3},\ \ldots.\gamma_{j(N-1)}]$$

$$j=1 \text{ to } 4 \quad (51)$$

P_{Ni} are polynomials in $\gamma_{j1}, \gamma_{j2}, \ldots.\gamma_{j(n-1)}$.

The domain of convergence may be estimated by the following inequality:

$$|\tau|<T'(1-\epsilon^{-\mu/5ST'}) \quad (52)$$

The constants S, μ, and T' are chosen so that $S\geq|g_i|,\mu_i>\mu$, and $T>T'$ for all values of y_i and τ, in a closed domain, defined by

$$|y_i(t+\tau)-y_i(t)|\leq\mu_i;\ |\tau|\leq T,\ i=1,2\ldots.4 \quad (53)$$

where $\mu_i>0$, $T>0$, and y_i's belong to a particular Ω_j, which implies that the system trajectories remain in the same region, Ω_j, in phase space during time interval $[t,\ t+\tau]$.

For short extrapolation time τ, equation 50 may be simplified by series truncation and by considering the oil volumes on each side of the actuator piston as constants during the extrapolation time interval $[t,\ t+\tau]$.

$$\left.\begin{aligned} V_1(\tau)&=[V_T/2+A_a x_L(\tau)]\approx[V_T/2+A_a x_L(t)] \\ V_2(\tau)&=[V_T/2-A_a x_L(\tau)]\approx[V_T/2-A_a x_L(t)] \end{aligned}\right\} \quad (54)$$

An estimate of truncation error is obtainable from the following inequality for the upper bound of the remainder R_k for a k-term truncated series solution:

$$|R_k|\leq\frac{\mu|\tau|^{k+1}}{(T'')^k(T''-|\tau|)}\times\left\{2+\left[\frac{5MT'}{\mu}\left|\ln\left(1-\frac{T''}{T'}\right)\right|\right]^{1/5}\right\} \quad (55)$$

where $|\tau|<T''<T'(1-\epsilon^{-\mu/5ST'})$ (56)

It should be remarked that many well-known numerical procedures for starting the solution of a system of differential equations (e.g., Runge-Kutta method) consist of matching the first few terms of the power series solutions. Moreover, in these procedures, no attempt is made to obtain solutions explicit in the independent variable τ, but rather to compute the change in the solution due to a given

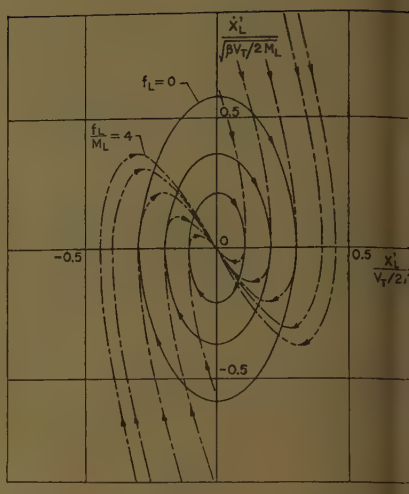


Fig. 5. Phase-plane trajectories of actua load after termination of a rectangular curr pulse $[P_1(T_D)=P_2(T_D)]$

increment in τ. The power series so tions, explicity expressed in terms of initial state variables, $y_i(t)$, and extrapo tion time τ, are useful in determining required forcing function to satisfy specified performance criterion. For ample, let the desired load position, $r(t+$ and velocity, $\dot{r}(t+\tau)$, within a finite ti interval, $(t, t+\tau)$, be specified in trunca power series form:

$$r(t+\tau)=r(t)+\dot{r}(t)\tau+\sum_{n=2}^{N}\Gamma_n(t)\tau^n$$

$$\dot{r}(t+\tau)=\dot{r}(t)+\sum_{n=2}^{N}n\Gamma_n(t)\tau^{n-1} \quad ($$

Then, a possible approach to the cont problem is to find the required port a coefficients, α_n, in equation 49 so tha prescribed performance criterion defi over the finite time interval, $[t,\ t+$ may be satisfied. A typical class of p formance criteria has the form:

$$\text{minimize}\int_0^{\tau}\phi[r(t+\tau')-x_L(t+\tau'),\ \dot{r}(t+\tau')-\dot{x}_L(t+\tau')]d\tau' \quad ($$

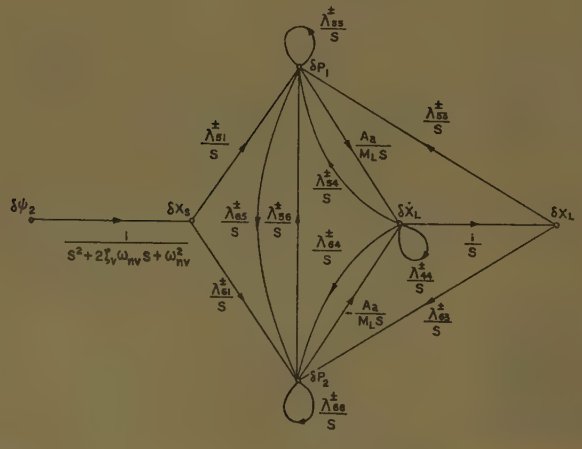


Fig. 6 (left). Flow graph for linearized system

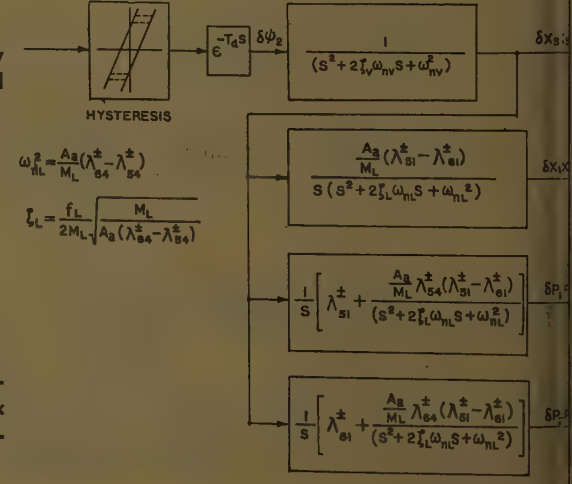


Fig. 7 (right). Transfer function block diagram for a simplified system

ere ϕ is a specified function of its argu-
nts.

ESPONSE TO A RECTANGULAR CURRENT PULSE

In time-optimal and pulse-width-modu-
ed sampled-data control systems, the
quired forcing of the output dynamic
ember often takes the form of a time
quence of positive and negative rectan-
lar pulses.[13] The valve-actuator re-
onse to such a current-pulse sequence
n be readily deduced from the subse-
ent discussion on single-pulse response.
First, consider the case where the valve-
ool motion is instantaneous with respect
torque-motor current. The valve ports
en upon pulse initiation at time, $\tau=0$,
d close upon pulse termination at $\tau=T_p$.
nce the system response during the
lse duration, $0<\tau<T_p$, is identical to the
ep response, only the trajectories after
rt closure remain to be determined.
learly, they correspond to the free-
cillation trajectories of actuator load
d the closed oil columns between the
tuator piston and the closed valve ports.
Setting $A_p[x_s(\tau)]$, in equations 19 and
), equal to zero and performing necessary
tegrations, the equation for load mo-
on becomes

$$\ddot{x}_L'(\tau')+\frac{F_L[\dot{x}_L'(\tau')]}{M_L}+\frac{A_a\beta}{M_L}\ln\left\{\left[1+\frac{x'_L(\tau')}{\frac{V_T}{2A_a}+x_L(T_p)}\right]\Big/\left[1-\frac{x_L'(\tau')}{\frac{V_T}{2A_a}-x_L(T_p)}\right]\right\}$$
$$=\frac{A_a}{M_L}[P_1(T_p)-P_2(T_p)] \quad (59)$$

here $\tau'=\tau-T_p$ and $x_L'(\tau'=0)=x_L(T_p)$.
he closed oil columns are in effect non-
near hard springs.
Equation 59 is valid if the pressures P_1
nd P_2 never drop down to the oil vapor
ressure P_v.
The first integral for equation 59 with
$_L=0$ is

$$\{[\dot{x}_L'(\tau')]^2-[\dot{x}_L(T_p)]^2\}$$
$$=\frac{A_a}{M_L}\{x_L'(\tau')[P_1(T_p)-P_2(T_p)]\}-$$
$$\frac{A_a\beta}{M_L}\ln\left\{\left(1+\frac{x_L'(\tau')}{\left[\frac{V_T}{2A_a}+x_L(T_p)\right]}\right)^{\left[\frac{V_T}{2A_a}+x_L(T_p)+x_L'(\tau')\right]}\times\right.$$
$$\left.\left(1-\frac{x_L'(\tau')}{\left[\frac{V_T}{2A_a}-x_L(T_p)\right]}\right)^{\left[\frac{V_T}{2A_a}-x_L(T_p)-x_L'(\tau')\right]}\right\} \quad (60)$$

For $F_L[\dot{x}_L'(\tau')]\neq 0$, the trajectories
ay be constructed in the phase plane
y using the isocline method. Typical
ajectories for damped and undamped
stems with $P_1(T_p)=P_2(T_p)$ are shown
Fig. 5. In practical systems, the single-
lse test is useful for experimental deter-
mination of total actuator load damping provided that the valve is quick-acting.

A more realistic case would be to describe the port areas closure by a truncated power series in τ'.

$$A_p[x_s(\tau')]=A_p(T_p)+\sum_{n=1}^{N}\alpha_n(\tau')^n \quad (61)$$

For a value with constant port width, W_o, and the spool motion describable by equation 3 with $F_T=0$, (since $i_M=0$ after pulse termination), equation 61 takes the form:

$$A_p[x_s(\tau')]=W_o\left\{x_s(T_p)+\dot{x}_s(T_p)\tau'-\right.$$
$$\left.[2\zeta_v\omega_{nv}\dot{x}_s(T_p)+\omega_{nv}^2x_s(T_p)]\frac{\tau'^2}{2}+\dots\right\} \quad (62)$$

Approximate expressions for load trajectories are obtainable by direct application of equations 50 and 51.

$$\dot{x}_L(\tau')=\dot{x}_L(T_p)+\sum_{n=1}^{\infty}\gamma_{2n}(\tau')^n \quad (63)$$

$$x_L(\tau')=x_L(T_p)+\dot{x}_L(T_p)\tau'+\sum_{n=1}^{\infty}\frac{\gamma_{2n}}{n}(\tau')^{n+1} \quad (64)$$

Here, τ' must be restricted to the non-cavitating range.

For the case where $F_L=0$, $\dot{x}_L(T_p)=C_vA_p(T_p)\sqrt{P_s/2}/A_a$, and $P_1(T_p)=P_2(T_p)=P_s/2$, the first few series coefficients are

$$\gamma_{21}=0$$
$$\gamma_{22}=0$$
$$\gamma_{23}=\frac{A_a\beta C_v\alpha_1}{6M_L}\sqrt{\frac{P_s}{2}}\left[\frac{1}{V_1(T_p)}+\frac{1}{V_2(T_p)}\right]$$

$$\gamma_{24}=\frac{A_a}{12M_L}\left\{\beta\alpha_2C_v\sqrt{\frac{P_s}{2}}\left[\frac{1}{V_1(T_p)}+\frac{1}{V_2(T_p)}\right]-\frac{\beta^2C_v^2\alpha_1A_p(T_p)}{4}\times\left[\frac{1}{V_1^2(T_p)}+\frac{1}{V_2^2(T_p)}\right]\right\} \quad (65)$$

It can be seen that if $\alpha_1=W_o\dot{x}_s(T_p)=0$, then γ_{23} is identically zero. Hence the series solution for $\dot{x}_L(\tau')$ starts with $\dot{x}(T_p)$ and a term of the order $(\tau')^4$. The case just cited is a rather special one. In other situations, the first three coefficients never vanish simultaneously.

The results derived in this section have the following applications:

1. They may be applied to near time-optimal electrohydraulic servomechanism design. Since the oil column between the closed-valve ports and the actuator piston provides an effective means of braking the load motion, a possible operating mode for achieving a near time-optimal noncavitating response (for step positioning) is to open the valve ports as wide as possible during the initial acceleration period and to follow this by a controlled closure of the valve ports. The valve port itself acts as a nonlinear damper. The approximate port closure time function may assume the form of equation 59. By specifying the desired load response during port closure in the form of a polynomial in time, the required coefficients α_n may be computed.

2. Calculation of system response to a time sequence of rectangular pulses may

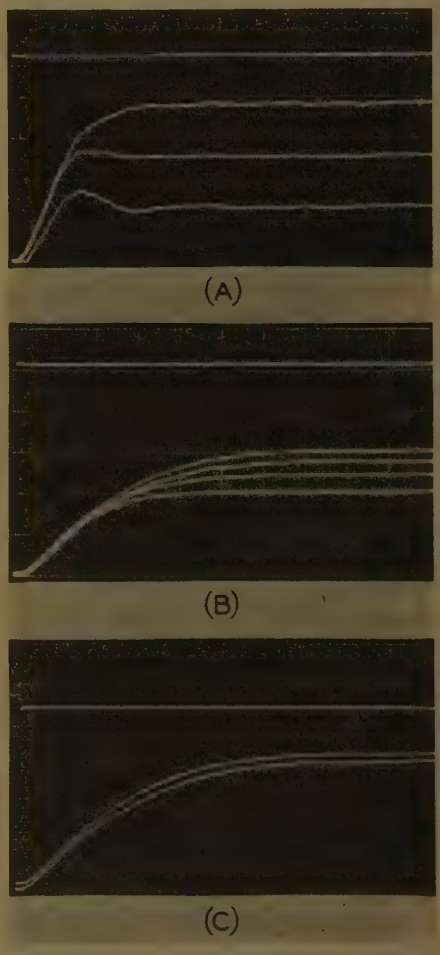


Fig. 8. Velocity response of the physical model to step-current inputs with a time scale of 5 milliseconds per large division and a velocity scale of 4.64 (A) and 11.6 inches per second per large division (B, C); current increments: 1 m

be pursued in a piece-wise manner using the results for single-pulse response.

Model for Small Motions

Linearization techniques are applicable for describing the small motion of nonlinear systems and investigating their local stability about their equilibrium states. Linearized models of valve-actuator systems have been discussed extensively in the past.[1-3,10] Here, a generalized linear nodel is derived systematically from the basic system equation given by 12. Due to discontinuities in the right-hand side of equation 12, the usual linearization procedure must be applied with caution, particularly in the case where the system's equilibrium state lies on one of the mode boundaries given by equation 25. In a precision servomechanism with a "zero-lapped" valve and step position inputs, a mathematical model linearized about $x_s(\tau)=0$ is useful for small-motion and stability studies of the closed-loop system. In the case where the input to the servomechanism is of ramp nature, the torque-motor operates with a bias current, i_{Mo}, and the steady-state load motion will have a velocity component. Then the linearized model is derived by considering the load position, $x_L(\tau)$, quasi-stationary in time.

Let the equilbrium and perturbed state vectors, denoted by $\mathbf{U}_e^{\pm}$ and $\delta\mathbf{U}^{\pm}$ respectively, be defined as

$$\mathbf{U}_e^{\pm}=\begin{bmatrix} x_{se}^{\pm} \\ \dot{x}_{se}^{\pm} \\ x_{Le}^{\pm} \\ \dot{x}_{Le}^{\pm} \\ P_1^{\pm} \\ P_2^{\pm} \end{bmatrix}, \quad \delta\mathbf{U}=\mathbf{U}-\mathbf{U}_e^{\pm}=\begin{bmatrix} \delta x_s \\ \delta\dot{x}_s \\ \delta x_L \\ \delta\dot{x}_L \\ \delta P_1 \\ \delta P_2 \end{bmatrix} \tag{66}$$

The $(\,)^+$ and $(\,)^-$ notations are used again to denote the system state corresponding to positive and negative x_{se} respectively. For $x_{se}=0$, the notations indicate the sign of δx_s.

Linearization of equation 12 about the equilibrium state leads to

$$\frac{d(\delta\mathbf{U}^+)}{d\tau}=\Lambda^+\delta\mathbf{U}^+ + \delta\mathbf{\Psi}(\tau) \tag{67}$$

and

$$\frac{d(\delta\mathbf{U}^-)}{d\tau}=\Lambda^-\delta\mathbf{U}^- + \delta\mathbf{\Psi}(\tau) \tag{68}$$

where Λ^+ and Λ^- are the Jacobian matrices, associated with $\mathbf{H}(\mathbf{U})$, evaluated at $\mathbf{U}_e^+$ and $\mathbf{U}_e^-$ respectively. Their explicit forms are

$$\Lambda^{\pm}=\begin{Vmatrix} 0 & 1 & 0 & 0 \\ -\omega_{nv}^2 & -2\zeta_v\omega_{nv} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \lambda_{44}^{\pm} \\ \lambda_{51}^{\pm} & 0 & \lambda_{53}^{\pm} & \lambda_{54}^{\pm} \\ \lambda_{61}^{\pm} & 0 & \lambda_{63}^{\pm} & \lambda_{64}^{\pm} \end{Vmatrix}$$

$$\lambda_{ij}^{+}=\left.\frac{\partial h_i^{+}}{\partial u_j}\right|_{\mathbf{U}_e^{+}}, \quad \lambda_{ij}^{-}=\left.\frac{\partial h_i^{-}}{\partial u_j}\right|_{\mathbf{U}_e^{-}}$$

$\delta\mathbf{\Psi}(\tau)$ is a perturbed forcing vector by

$$\delta\mathbf{\Psi}(\tau)=\begin{bmatrix} 0 \\ \delta\psi_2(\tau) \\ 0 \\ \delta\psi_4(\tau) \\ 0 \\ 0 \end{bmatrix}=\mathbf{\Psi}(\tau)-\begin{bmatrix} 0 \\ F_T(i_{Mo})/M_s \\ 0 \\ F_{Do}/M_L \\ 0 \\ 0 \end{bmatrix} \tag{70}$$

where F_{Do} is the mean value of the external load disturbance force.

Performing a Laplace transformation of equations 67 and 68 leads to

$$\begin{aligned} (sI-\Lambda^+)\delta\mathbf{U}^+(s)&=\delta\mathbf{\Psi}(s) \\ (sI-\Lambda^-)\delta\mathbf{U}^-(s)&=\delta\mathbf{\Psi}(s) \end{aligned} \tag{71}$$

Equivalent transfer functions relating $\delta\psi_2(s)$ and $\delta\psi_4(s)$ with various components of $\delta\mathbf{U}^{\pm}$ are obtainable by reduction of the flow graph of Fig. 6 associated with equation 71.

For a symmetrical, "zero-lapped," rectangular-port (width W_o) valve with viscous load friction and negligible leakage, the equilibrium state vector for $i_{Mo}=0, (x_{se}=0)$ is

$$\mathbf{U}_e=\begin{bmatrix} 0 \\ 0 \\ x_{Le} \\ 0 \\ P_s/2 \\ P_s/2 \end{bmatrix} \tag{72}$$

The matrix elements $\lambda_{ij}^{\pm}$ are

$$\left.\begin{aligned} \lambda_{44}^{\pm}&=-\frac{f_L}{M_L} \\ \lambda_{51}^{\pm}&=\frac{\beta C_v W_o}{\left(\frac{V_T}{2}+A_a x_{Le}\right)}\sqrt{\frac{P_s}{2}} \\ \lambda_{61}^{\pm}&=\frac{-\beta C_v W_o}{\left(\frac{V_T}{2}-A_a x_{Le}\right)}\sqrt{\frac{P_s}{2}} \\ \lambda_{54}^{\pm}&=\frac{-\beta A_a}{\left(\frac{V_T}{2}+A_a x_{Le}\right)} \\ \lambda_{64}^{\pm}&=\frac{\beta A_a}{\left(\frac{V_T}{2}-A_a x_{Le}\right)} \\ \lambda_{55}^{\pm}&=\lambda_{56}^{\pm}=\lambda_{65}^{\pm}=\lambda_{66}^{\pm} \\ &=\lambda_{53}^{\pm}=\lambda_{63}^{\pm}=0 \end{aligned}\right\} \tag{7}$$

As a result of valve symmetry, th corresponding matrix elements for $\delta x_s>$

$$\begin{Vmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{A_a}{M_L} & -\frac{A_a}{M_L} \\ \lambda_{55}^{\pm} & \lambda_{56}^{\pm} \\ \lambda_{65}^{\pm} & \lambda_{66}^{\pm} \end{Vmatrix} \tag{6}$$

and $\delta x_s<0$ are identical, so that the sy tem's small motion about $x_s=0$ can described by a single vector equatio For a value with nonsymmetrical fl characteristics, two equations in t forms given by 71 are necessary f complete description.

A transfer function diagram is shown Fig. 7. It can be seen from the resu that the damping of the actuator pob is governed by load friction force alon This is expected, since for small oil flo the damping effects of the valve ports a negligible. Valve underlap and leaka across the actuator piston generally i troduce additional damping. Hystere and dead-time in the valve torque-mot are included for stability analysis. He the well-known describing function tec niques may be used.

Experimental Studies

To check the accuracy of the deriv mathematical models, experimental te were performed on a typical valve-co trolled actuator system and its cor sponding mathematical model simulat by an analog computer. Pertine system parameters are listed in Appe dix II.

First, the step-current response of t system was determined. Fig. 8 sho

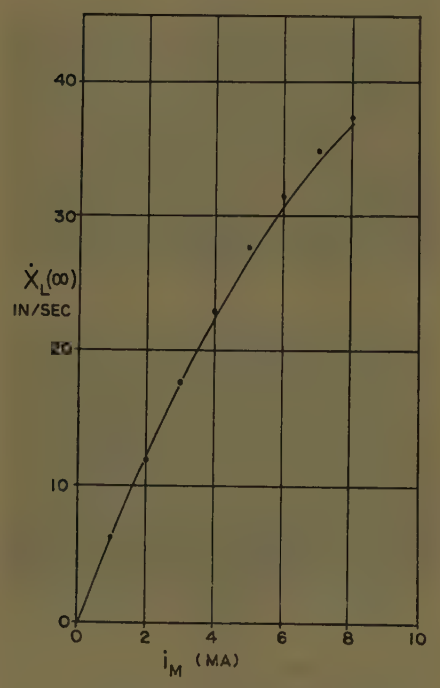


Fig. 9. Experimental and calculated steady-state velocities of the physical model with step-current inputs ($f_L=1.6$ pounds per inch per second)

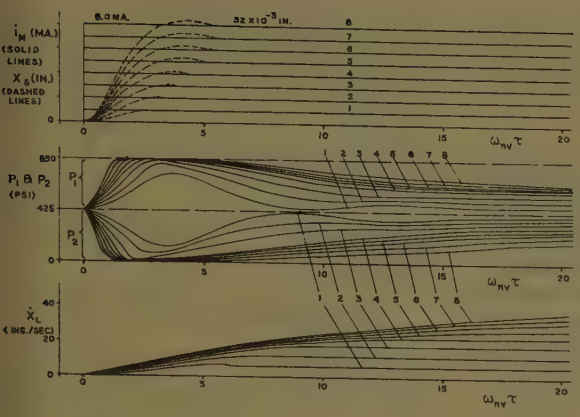


Fig. 10. Analog computer solutions of system's step-current response ($\omega_{nv}=943$ radians per second)

The velocity responses for various step-current amplitudes. Fig. 9 shows that the steady-state velocities are in close agreement with the values computed using equation 32. The viscous friction coefficient f_L used in the above calculations was experimentally estimated from single rectangular pulse test results. The velocity profiles for large currents agree within $\pm 5\%$ with the corresponding analog computer solutions in Fig. 10 and the approximate analytical result given by equation 42. For small current amplitudes, less than 2.0 ma (milliamperes), the velocity responses of the physical model become less damped as a result of the weaker damping effect of the valve ports, whereas this effect is predominant in large current response.

Figs. 11 and 12 show the system responses for single- and multiple-pulse inputs. Actuator pressures P_1 and P_2 are monitored by means of pressure transducers. The multiple mode nature of the system is clearly revealed in the lower portion of Fig. 12. The difference in rise times is shown by the velocity responses to step currents, +2 and +4 ma, which agree with the approximate results predicted by equation 42. After the current has returned to zero, the oscillatory response is due to the motion of the actuator load coupled with the closed oil columns between the valve port and actuator piston. High-pressure peaks exceeding P_s and cavitation are clearly indicated.

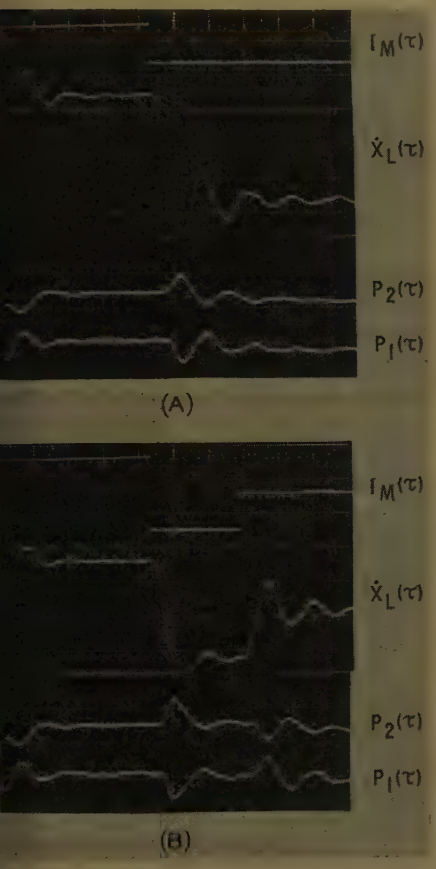


Fig. 11. (A) Single current-pulse response of the physical model with a pulse amplitude of 1 ma and a velocity scale of 2.32 inches per sec per large division. (B) Double current-pulse response of the physical model with a pulse amplitude of +1 ma and a velocity scale of 4.64 inches per second per large division. Time scale: 10 milliseconds per large division; pressure scale: 708 pounds per square inch per large division

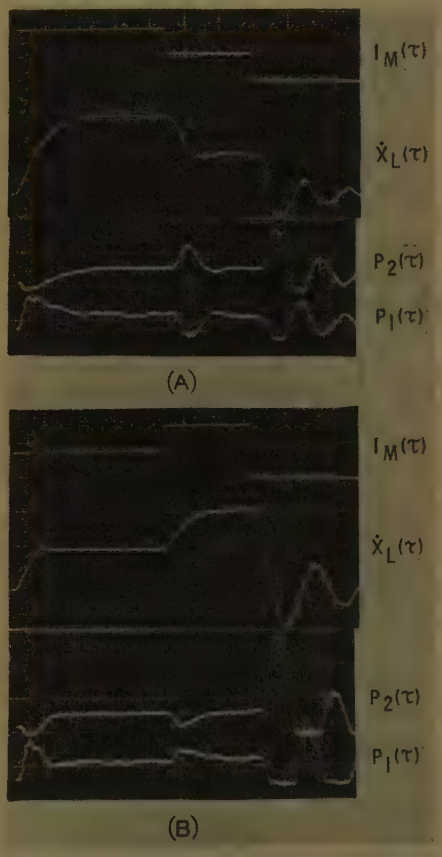


Fig. 12. Response of the physical model to a sequence of current pulses with a velocity scale of 11.6 inches per second per large division, a pressure scale of 708 pounds per square inch per large division, and pulse sequences of +4, +2, 0 ma (A) and +2, +4, 0 ma (B)

Conclusions

It has been shown that an electro-hydraulic valve controlled actuator can be represented by reasonably simple and accurate mathematical models by making suitable assumptions. The basic model is shown to be a multiple-mode nonlinear system. Simple results are derived for step and pulsed valve motor-current inputs and their applications are discussed. Finally, the adequacy of the models for time-domain design is substantiated by experimental studies of a physical model and its corresponding mathematical model simulated by an analog computer.

Appendix I. h_5 and h_6 for Underlapped and Overlapped Valves

Underlapped Valve (With Lapping $\pm\Delta_u$)

For $|\ X_s\ (T)\ | < \Delta_u$

$$h_5=\frac{\beta}{\left[\frac{V_T}{2}+A_a x_L(\tau)\right]}\{C_v'|A_p\, x_s(\tau)]|\times [\text{sgn}\,[P_s-P_1(\tau)]\sqrt{|P_s-P_1(\tau)|}-\text{sgn}\,[P_1(\tau)-P_r]\sqrt{|P_1(\tau)-P_r|}]-A_a\dot{x}_L(\tau)+C_L[P_1(\tau)-P_2(\tau)]\} \quad (74)$$

$$h_6=\frac{\beta}{\left[\frac{V_T}{2}-A_a x_L(\tau)\right]}\{C_v'|A_p[x_s(\tau)]|\times [\text{sgn}\,[P_s-P_2(\tau)]\sqrt{|P_s-P_2(\tau)|}-\text{sgn}\,[P_2(\tau)-P_r]\sqrt{|P_2(\tau)-P_r|}]+A_a\dot{x}_L(\tau)-C_L[P_1(\tau)-P_2(\tau)]\} \quad (75)$$

C_v' is a discharge proportionality constant for small port opening.

For $x_s(\tau)>\Delta_u$: same as equations 19 and 20

For $x_s(\tau)<-\Delta_u$: same as equations 21 and 22

Overlapped Valve (With Lapping $\pm\Delta_o$)

For $|x_s(\tau)|<\Delta_o$

$$h_5=\frac{-\beta A_a\dot{x}_L(\tau)}{\left[\frac{V_T}{2}+A_a x_L(\tau)\right]} \quad (76)$$

$$h_6=\frac{\beta A_a\dot{x}_L(\tau)}{\left[\frac{V_T}{2}-A_a x_L(\tau)\right]} \quad (77)$$

For $x_s(\tau)>\Delta_o$

$$h_5=\frac{\beta}{\left[\frac{V_T}{2}+A_a x_L(\tau)\right]}\{C_v A_p[x_s(\tau)-\Delta_o]\times \operatorname{sgn}[P_s-P_1(\tau)]\sqrt{|P_s-P_1(\tau)|}-A_a\dot{x}_L(\tau)+C_L[P_1(\tau)-P_2(\tau)]\} \quad (78)$$

$$h_6=\frac{\beta}{\left[\frac{V_T}{2}-A_a x_L(\tau)\right]}\{-C_v A_p[x_s(\tau)-\Delta_o]\times \operatorname{sgn}[P_2(\tau)-P_r]\sqrt{|P_2(\tau)-P_r|}+A_a\dot{x}_L(\tau)-C_L[P_1(\tau)-P_2(\tau)]\} \quad (79)$$

For $x_s(\tau)<-\Delta_o$

$$h_5=\frac{\beta}{\left[\frac{V_T}{2}+A_a x_L(\tau)\right]}\{-C_v|A_p[x_s(\tau)+\Delta_o]|\times \operatorname{sgn}[P_1(\tau)-P_r]\sqrt{|P_1(\tau)-P_r|}-A_a\dot{x}_L(\tau)+C_L[P_1(\tau)-P_2(\tau)]\} \quad (80)$$

$$h_6=\frac{\beta}{\left[\frac{V_T}{2}-A_a x_L(\tau)\right]}\{C_v|A_p[x_s(\tau)+\Delta_o]|\times \operatorname{sgn}[P_s-P_2(\tau)]\sqrt{|P_s-P_2(\tau)|}+A_a\dot{x}_L(\tau)-C_L[P_1(\tau)-P_2(\tau)]\} \quad (81)$$

Appendix II. Parameters of the Physical Model

Actuator

Inertial load (M_L): 0.0311 pound second squared per inch (12-pound weight)
Load Coulomb friction force (F_c): approximately 4 pounds
Load viscous friction force coefficient f_L, determined from single-pulse response: 1.6 pounds per inch per second
Piston area (A_a): 0.125 inch squared
Maximum piston stroke: 5.0 inches
Total volume of oil in cylinder (V_T): 0.99 inch^3
Bulk modulus of oil (β): 2.5×10^5 pounds per square inch

Valve (2-Stage Flapper Type)

$\omega_{nv}=943$ radians per second (150 cycles per second)
$\zeta_v=0.7$
Port width (W_o): 0.1 inch
Spool displacement for $i_M=4.0$ ma: 0.016 inch
Port discharge proportionality constant: 104 inches^2/(pound)$^{1/2}$ second

Hydraulic Supply

Supply pressure (P_s): 850 pounds per square inch (gage)
Return pressure (P_r): atmospheric

References

1. Fluid Power Control (book), J. F. Blackburn, G. Reethof, J. L. Shearer, editors. John Wiley & Sons, Inc., New York, N. Y., 1960.

2. Dynamic Operation of a Force-Compensated Hydraulic Throttling Valve, J. L. Bower and F. B. Tuteur. *Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 75, 1953, pp. 1395–1406.

3. An Analysis of the Dynamics of Hydraulic Servomotors Under Inertia Loads and the Application to Design, H. Gold, E. W. Otto, V. L. Ransom. *Ibid.*, pp. 1383–94.

4. A Theoretical Analysis of the Response of a Loaded Hydraulic Relay, R. Butler. *Proceedings*, Institute of Mechanical Engineers, London, England, vol. 173, no. 16, 1959, pp. 429–58.

5. The Response of a Loaded Hydraulic Servomechanism, D. E. Turabull. *Ibid.*, no. 9, 1959, pp. 270–93.

6. A Describing Function for Multiple Nonlinearities Present in Electrohydraulic Control Valves, J. Zaborszky, H. J. Harrington. *AIEE Transactions*, pt. I (*Communication and Electronics*), vol. 76, May 1957, pp. 183–90.

7. Generalized Charts of the Effects of Nonlinearities in Electrohydraulic Control Valves, J. Zaborszky, H. J. Harrington. *Ibid.*, pp. 191–98.

8. A Describing Function for the Multiple Nonlinearities Present in 2-Stage Electrohydraulic Control Valves, J. Zaborszky, H. J. Harrington. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 76, Jan. 1958, pp. 394–01.

9. Generalized Charts of the Effects of Nonlinearities in 2-Stage Electrohydraulic Control Valves, J. Zaborszky, H. J. Harrington. *Ibid.*, pp. 401–08.

10. The Analysis of Valve-Controlled Hydraulic Servomechanisms, R. G. Rausch. *Bell System Technical Journal*, New York, N. Y., vol. 38, 1959, pp. 1513–60.

11. Contributions to Hydraulic Control, Part I and II, S. Y. Lee, J. F. Blackburn. *Transactions*, American Society of Mechanical Engineers, vol. 74, 1952, pp. 1005–16.

12. Relay-Type Control Systems Designed for Random Inputs, A. M. Hopkin, P. K. C. Wang. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 78, Sept. 1959, pp. 227–33.

13. Pulse Width Control of Sampled-Data Systems, W. L. Nelson. *Technical Report 35/B*, Dept. of Elec. Eng., Columbia University, New York, N. Y., July 1959.

Signal Stabilization of a Control System with Random Inputs

R. OLDENBURGER
MEMBER AIEE

R. SRIDHAR
NONMEMBER AIEE

NOISE AND ITS EFFECT on system performance now can be predetermined by a designer if he uses the theory advanced herein. An external signal is introduced to stabilize a nonlinear feedback system in a state of self-sustained oscillation, but the introduction should be made at a sufficiently high frequency and at a convenient point in the loop. This phenomenon, first introduced by R. Oldenburger, one of the authors, was termed "signal stabilization."[1]

This author, in collaboration with Liu,[2] later explained signal stabilization in the case of a system with one nonlinear element when the external or stabilizing signal was sinusoidal. They defined an equivalent gain for a nonlinearity and showed that the latter can often be replaced by the former for stability analysis. Equivalent gain was described as the "limit of the output divided by the average value of the input ..." as the latter value "goes to zero." The input was, at that time, assumed to be a biased sine wave. The theory developed by Oldenburger was unlike that published by Minorsky regarding the asynchronous quenching of systems described by nonlinear differential equations.[3]

Paper 61-712, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE-AIChE-ASME-IRE-ISA Joint Automatic Control Conference, Boulder, Colo., June 28–30, 1961. Manuscript submitted October 5, 1960; made available for printing May 8, 1961.

R. Oldenburger and R. Sridhar are with Purdue University, Lafayette, Ind.

Research forming the basis of this paper was supported by a grant from the National Science Foundation of the United States Government.

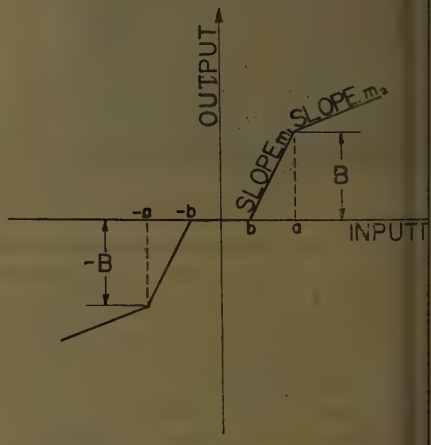


Fig. 1. Input-output characteristic, type nonlinearity

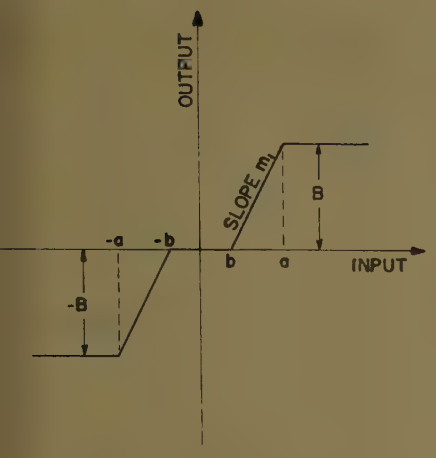


Fig. 2. Input-output characteristic, type B nonlinearity (limiter with dead band)

Oldenburger's signal-stabilization theory was extended, in collaboration with Nakada,[4] to self-oscillating systems; it is here further extended, in collaboration with Sridhar, to the case of a nonlinear system, which has one single-valued nonlinearity in the loop when the stabilizing signal belongs to a stationary random process with a Gaussian distribution and obeys the ergodic hypothesis. Now the definition of equivalent gain takes on wider scope than originally stated, and thereby becomes much more useful. The original actually turns out to be a particular case within the more general meaning of the new definition. Using this general precept, a method was found and is here presented for determining the mean-square value of noise to be injected at the input to certain self-oscillating nonlinear systems in order to reduce the output hunt to a desired value.

Derivation of Equivalent Gains

Only nonmemory-type nonlinearities,[5] whose output y may be represented as a single-valued odd function $f(x)$ of the input x, are within the restrictions set for this paper. The two general types to be discussed are those nonlinearities (1) whose input-output characteristic consists of straight-line segments, designated as "piecewise-linear," and (2) whose input-output characteristic may be described by a polynomial, designated as "polynomial."

The equivalent gain $g(m)$ of a nonmemory-type of nonlinearity is the ratio between the average value A_y of the output y and the average value m of the input x, when the input to the nonlinearity consists of a stationary random process with a Gaussian distribution. The first probability density function $p_1(x)$ of the input is given by

$$p_1(x)=\frac{1}{\sqrt{2\pi\phi_0}}\exp\left[-\frac{(x-m)^2}{2\phi_0}\right] \quad (1)$$

where ϕ_0 is the variance of the input. Thus

$$g(m)=\frac{A_y}{m} \quad (2)$$

The limiting value of the equivalent gain defined by equation 2, when $m\to 0$, reduces to the equivalent gain defined by Oldenburger and Liu. This limiting value will be denoted as $g(0)$.

Equivalent gains of the two general types of nonlinearities will be derived next.

Piecewise-Linear-Type Nonlinearity

The different piecewise-linear-type nonlinearities are shown in Figs. 1 through 7, and their input-output characteristics are described in Appendix I. Fig. 1 represents the most general type and the other six are special cases of this nonlinearity. They will be classified alphabetically from A to G for the sake of clarity. Thus, Fig. 1 corresponds to A, Fig. 2 to B, and so on. To distinguish between expressions for the equivalent gains of each nonlinearity, classifying letters of the alphabet will be added as subscripts to $g(m)$. Thus $g_a(m)$ corresponds to the equivalent gain of the type-A nonlinearity, while $g_a(0)$ corresponds to the gain of the same type when evaluated at $m=0$.

The only expression for equivalent gain which must be derived is that of the type-A nonlinearity. Expressions for other types are obtained by considering them as special cases of the type-A nonlinearity.

The general expression for A_y is

$$A_y=\int_{-\infty}^{\infty} f(x)p_1(x)dx \quad (3)$$

The expression for A_y, evaluated in Appendix I for the type-A nonlinearity, yields

$$g_a(m)=\frac{m_1-m_2}{m}\sqrt{\frac{\phi_0}{2\pi}}\left[\exp\left\{-\frac{(a+m)^2}{2\phi_0}\right\}-\exp\left\{-\frac{(a-m)^2}{2\phi_0}\right\}\right]+m_2-\frac{m_1}{m}\sqrt{\frac{\phi_0}{2\pi}}\left[\exp\left\{-\frac{(b+m)^2}{2\phi_0}\right\}-\exp\left\{-\frac{(b-m)^2}{2\phi_0}\right\}\right]+\frac{m_1-m_2}{2}\times\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)+\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right]+\frac{m_1-m_2}{m}\times\frac{a}{2}\times\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right]-\frac{m_1}{2}\left[\operatorname{erf}\left(\frac{b+m}{\sqrt{2\phi_0}}\right)+\operatorname{erf}\left(\frac{b-m}{\sqrt{2\phi_0}}\right)\right]-\frac{m_1}{m}\times\frac{b}{2}\times\left[\operatorname{erf}\left(\frac{b+m}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{b-m}{\sqrt{2\phi_0}}\right)\right] \quad (4)$$

Taking the limiting value of the right side of equation 4 and simplifying the resulting expression gives

$$g_a(0)=m_2\left[1-\operatorname{erf}\left(\frac{a}{\sqrt{2\phi_0}}\right)\right]+m_1\left[\operatorname{erf}\left(\frac{a}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{b}{\sqrt{2\phi_0}}\right)\right] \quad (5)$$

Expressions for equivalent gains of the other nonlinearities are now easily derived from equations 4 and 5, as illustrated in Appendix II.

Equivalent gains of the type-C nonlinearity (limiter) and the type-G nonlinearity (ideal relay) are shown in Figs. 8 and 9, respectively.

Polynomial Type of Nonlinearity

The general polynominal type of nonlinearity is described by equation 6.

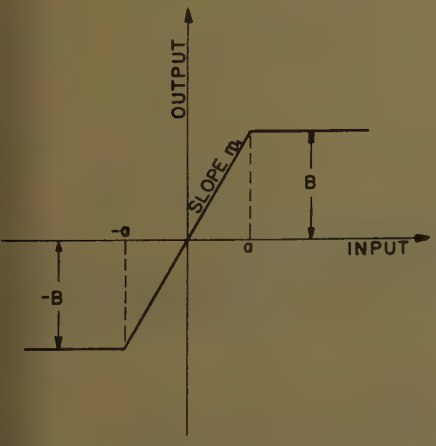


Fig. 3. Input-output characteristic, type C nonlinearity (limiter)

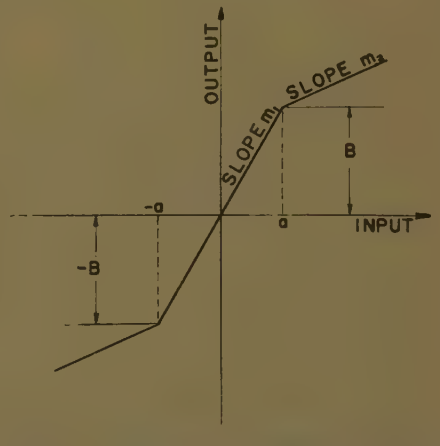


Fig. 4. Input-output characteristic, type D nonlinearity

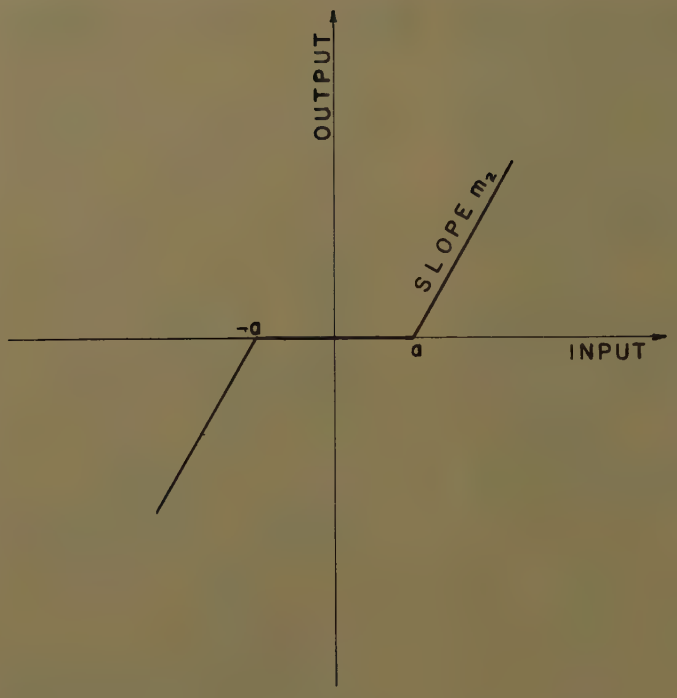


Fig. 5. Input-output characteristic, type E nonlinearity (dead-zone element)

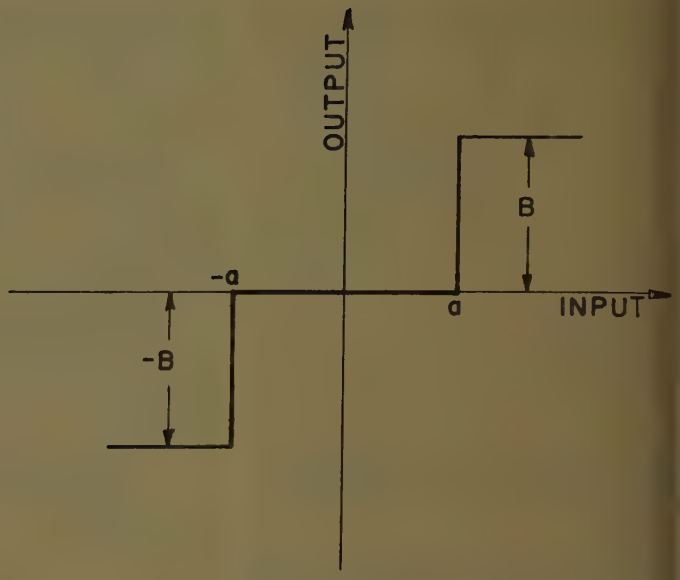


Fig. 6. Input-output characteristic, type F nonlinearity (relay with dead band)

There is no loss of generality in assuming that n is odd in

$$y=f(x)=c_nx^n+c_{n-1}x^{n-2}|x|+c_{n-2}x^{n-2}+\ldots c_{n-3}x^{n-4}|x|+\ldots+c_2x|x|+c_1x \quad (6)$$

The absolute-value signs have been used in the right side of equation 6 to make $f(x)$ an odd function of x.

In deriving the equivalent gain, each term of the polynomial is considered separately. Thus, for purpose of derivation, two types of nonlinearities are assumed. There is no loss in generality by this assumption since every term in the polynomial belongs to one of the two types shown in equations 7 and 8.

$$y=f(x)=cx^n \qquad \text{for } n \text{ odd} \qquad (7)$$

$$y=f(x)=\begin{cases}+cx^n & x\geq 0\\ -cx^n & x<0\end{cases} \qquad \text{for } n \text{ even} \qquad (8)$$

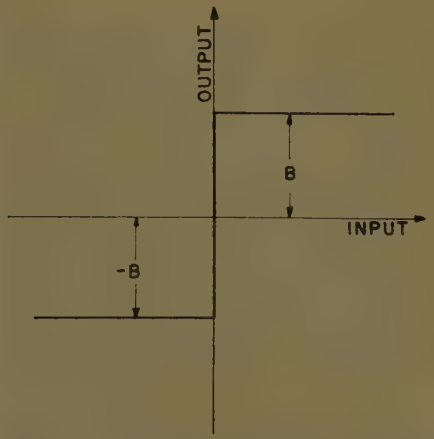


Fig. 7. Input-output characteristic, type G nonlinearity (ideal relay)

The derivation in Appendix III shows that the nonlinearity corresponding to equation 7 has an equivalent gain $g_1(m)$ of the form

$$g_1(m)=\frac{A_v}{m}=\frac{cm^{n-1}}{\sqrt{\pi}}\sum_{k=0,2,4}^{n-1}\times \frac{n!}{(n-k)!}\frac{1}{k!}\left(\frac{2\phi_0}{m}\right)^k\Gamma\left(\frac{k+1}{2}\right) \quad (9)$$

Taking the limiting value of the right side of equation 9 when $m\to 0$ yields

$$g_1(0)=\frac{c}{\sqrt{\pi}}n(\sqrt{2\phi_0})^{n-1}\Gamma\left(\frac{n}{2}\right) \quad (10)$$

Also in Appendix III, the equivalent gain $g_2(m)$ of the nonlinearity defined by equation 8 is given by

$$g_2(m)=\frac{A_v}{m}=\frac{c}{\sqrt{\pi}}m^{n-1}\left[\sum_{0,2,4}^{n}\frac{n!}{(n-k)!}\times \frac{1}{k!}\times\left(\frac{\sqrt{2\phi_0}}{m}\right)^k\times 2\int_0^{\frac{m}{\sqrt{2\phi_0}}}z^k\times \exp(-z^2)dz+\sum_{k=1,3,5}^{n-1}\frac{n!}{(n-k)!}\times \frac{1}{k!}\left(\frac{\sqrt{2\phi_0}}{m}\right)^k\times 2\int_{\frac{m}{\sqrt{2\phi_0}}}^{\infty}\times z^k\exp(-z^2)dz\right] \quad (11)$$

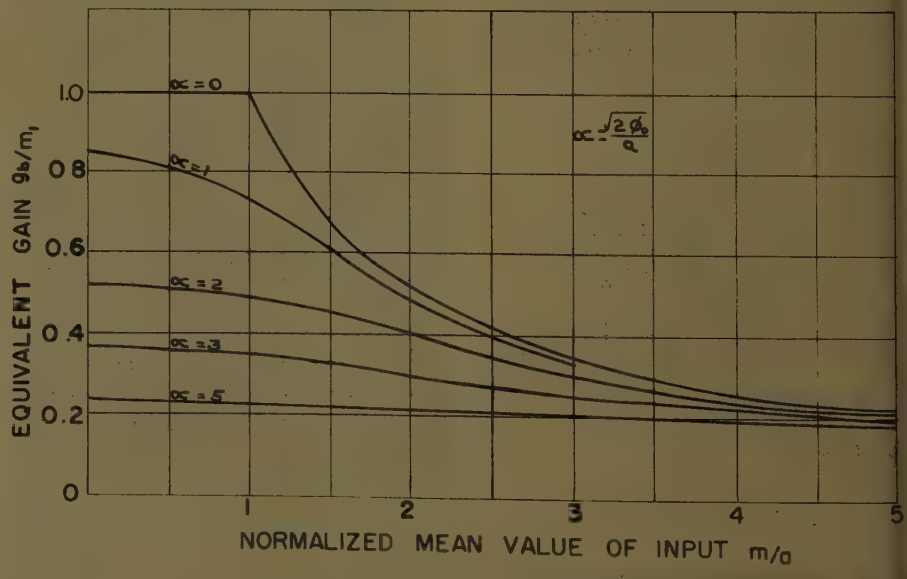


Fig. 8. Equivalent gain of limiter

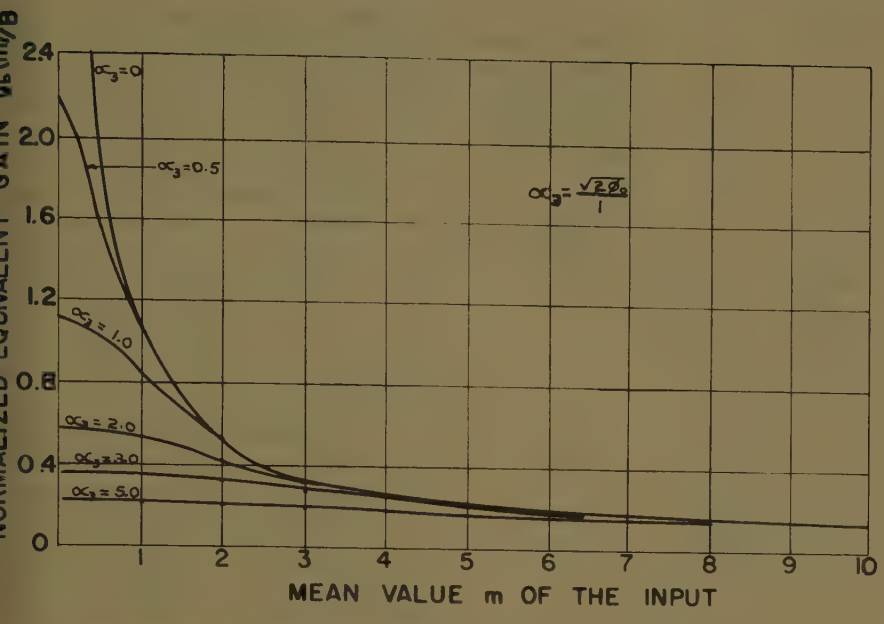


Fig. 9. Equivalent gain of ideal relay

Taking the limiting value of the right de of equation 11, when $m \to 0$, yields

$$(0)=\frac{2c}{\sqrt{\pi}}\times n(\sqrt{2\phi_0})^{n-1}\int_0^\infty \times$$

$$z^{n-1}\exp(-z^2)dz \quad (12)$$

The equivalent gain of the general olynomial-type nonlinearity, defined by uation 6, will equal the sum of all terms aving the form seen in the right sides of uations 9 and 11.

The particular case when n equals zero equation 8 results in the input-output aracteristic of the ideal relay (type-G onlinearity). Equating n to zero in the ght side of equation 11 should yield the uivalent gain of the ideal relay which as derived in Appendix I. Thus

$$g_2(m)=\frac{c}{m}\ \text{erf}\left(\frac{m}{\sqrt{2\phi_0}}\right)\ \text{for}\ n=0 \quad (13)$$

The right sides of equations 13 and 30 are—as they should be—the same.

Equivalent Admittance of a Nonlinearity

Before the theory of signal stabilization with random inputs is used in practical systems, it is advisable to understand "equivalent admittance" of a nonlinearity and also know how it may be determined.

Definition: The equivalent admittance $Jnl(A, \phi_0)$ of a nonlinearity is the complex ratio of the amplitude P of the output component, which has the same frequency as the input sine wave to the amplitude A of the input sine wave. (In this definition, it is assumed that the input consists of a sine wave, together with Gaussian noise, and that the power spectral density of the noise has no predominant component with the same frequency as the input sine wave.)

Table I. Fundamental and Third-Harmonic Components of Output Waveforms Shown in Figs. 10 and 11

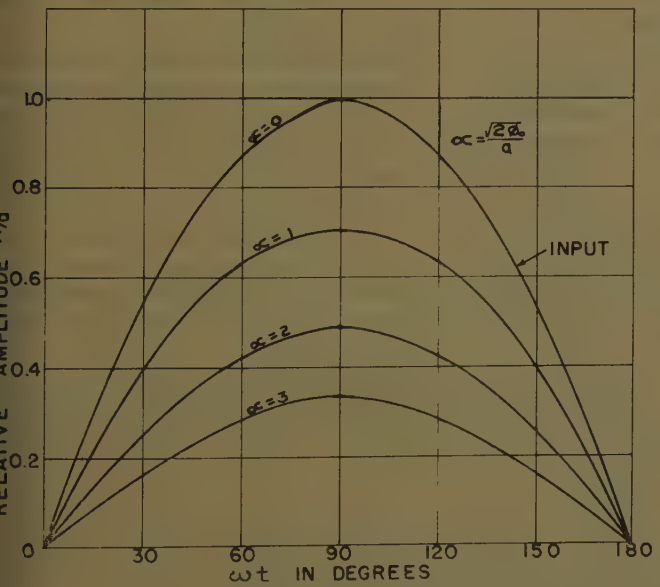
$\frac{\sqrt{2\phi_0}}{a}$ (Normalized Rms Value of Noise)	0	1	2	3
When A/a (Normalized Amplitude of Sine Wave) equals 1				
P_1/m_1a	1	0.745	0.507	0.3435
P_3/m_1a	0	0.0323	0.01165	0.00445
When A/a (Normalized Amplitude of Sine Wave) equals 2				
P_1/m_1a	1.224	1.092	0.861	0.652
P_3/m_1a	0	0.149	0.0538	0.0187

Notice the similarity between this definition and the describing function of a nonlinearity, which is $Jnl(A, 0)$.[6] For the types of nonlinearities considered in this paper, the equivalent admittance is a real quantity.

Determination of Equivalent Admittance

Knowledge must be acquired regarding transmission of a sine wave through a nonlinearity in the presence of noise before equivalent admittance can be determined. The concept of equivalent gain may be used, as subsequently explained, to determine the amplitude of the output component of the nonlinearity having the same frequency as the input sine wave.

If the input to a nonlinearity consists



g. 10. Waveform, output of limiter for sinusoidal input A sin ωt (only half of period shown)

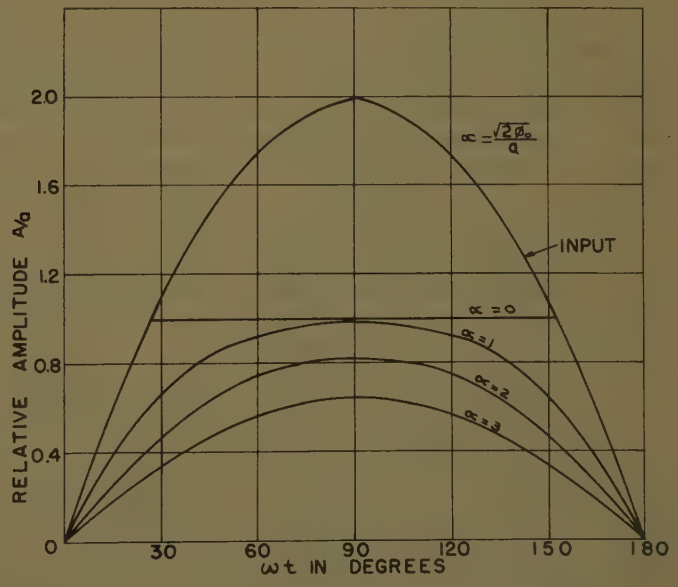


Fig. 11. Same as Fig. 10 for $\frac{A}{a}=2$ (only half of period shown)

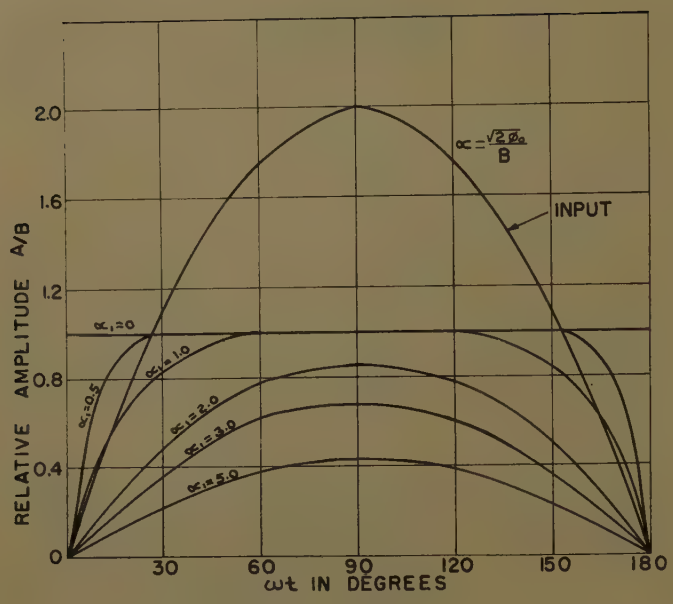


Fig. 12 Waveform, output of ideal relay for sinusoidal input A sin ωt (only half of period shown)

of zero-mean noise, together with an additional signal of constant value, then the gain or amplification of the d-c component of the input is the equivalent gain. No restrictions are placed on the frequency components of the noise. Now assume that the additional signal is varying instead of constant. If the lowest frequency component of the spectral density of the noise is appreciably higher than the highest frequency component of the additional signal, a reasonable conclusion is that the additional signal is almost constant compared with the noise. The validity of this assumption increases with an increase in the lowest frequency component of the spectral density of the noise.

Hence, if the input to a nonlinearity consists of a stabilizing signal and an additional signal, the latter varying slowly in comparison with the former, then the instantaneous gain of the additional signal approximately equals the equivalent gain.

This concept will be used now to determine the gain of a sinusoidal signal through a nonlinearity when noise is present.

Let this noise be stationary and have Gaussian distribution with a mean of zero and a mean-square value ϕ_0 (note that the mean-square value is equal to the variance in this case), and let the sine wave be of amplitude A and frequency ω. Consider half of one period of the sine wave as shown in Fig. 10. Divide one period into $2n$ intervals, which is the same as dividing half a period into n equal intervals, and erect ordinates at mid-points of each interval. When using the equivalent-gain technique, assume that the sine function is a constant, equal to the mid-ordinate value over each interval. This is essentially equivalent to approximating the sine wave by a series of steps.

Consider the kth interval. The abscissa of its middle point is $A \sin \pi/n(k-1/2) = k_1$. Since the input is assumed to be a constant equal to k_1 over the kth interval, the output value of the nonlinearity over this interval is $k_1\, g(k_1)$, where $g(k_1)$ is the equivalent gain of the nonlinearity for an argument k_1. The value of the output over each $2n$ interval may be determined similarly and the output waveform plotted. A series of steps comprise the output response. Fundamental and higher components of the output waveform may be determined now, using any of the well-known graphical or numerical methods.

Since the system beyond the nonlinearity is assumed to be essentially a low-pass filter, the components of the noise at the output may be neglected.

How to apply the equivalent-admittance method will be illustrated next by showing the transmission of a sine wave in the presence of noise through a limiter and a relay.

Figs. 10 and 11 show the input sine wave and the output waveforms for a limiter when A/a (normalized amplitude of the sine wave) is 1 and 2, respectively, and $\sqrt{2\phi_0}/a$ (normalized rms value of the noise) is 0, 1, 2, and 3. For convenience, a smooth curve has been drawn through the computed mid-ordinate points on the output waveform. One period of the sine wave has been divided into 36 equal parts for this computation.

The fundamental and third-harmonic components of the output waveforms in Figs. 10 and 11 were determined by using a 36-ordinate scheme over one full period, and the results are summarized in Table I, where P_1 and P_3 refer to the fundamental and third-harmonic components, respectively.

Similarly, Fig. 12 shows the input sine wave and the output waveforms for an ideal relay when A/B (normalized input amplitude) is 2 and $\sqrt{2\phi_0}/B$ is 0.5, 1, 2, 3, and 5. Fundamental and third-harmonic components of the output waveforms are shown in Table II. These components may be calculated in the same way for different nonlinearities by using knowledge of the equivalent gain of the nonlinearity. The ratio of this fundamental component amplitude to the input sine wave amplitude is the equivalent admittance. Equivalent-admittance curves for the limiter and the relay are shown in Figs. 13 and 14.

Reduction of System Hunting to a Desired Value

Signal stabilization is generally effective in a system having nonlinearities whose equivalent admittance $J_{s1}(A,\phi_0)$ is monotonically nondecreasing in both A and ϕ_0. The system should be absolutely stable below a particular value of gain when the nonlinearity is replaced by a simple gain, which means that the linear part of the system should not be conditionally stable.

In the discussion, it is assumed that the linear part of the system following the nonlinearity is such a good low-pass filter that noise feedback may be neglected.

In Fig. 15, $G_1(s)$, $G_2(s)$, and $H(s)$ represent the linear elements in the system and NL is the nonlinear element. The product $G_1(s) \times G_2(s) \times H(s)$ is the linear part of the loop and is designated $G(s)$

Table II. Fundamental and Third-Harmonic Components of Output Waveform Shown in Fig. 1

$\frac{\sqrt{2\phi_0}}{B}$ (Normalized Rms Value of Noise)	0	0.5	1	2	3	5
When A/B (Normal Input Amplitude) equals 2						
P_1/B	1.275	1.252	1.17	0.895	0.623	0.4
P_3/B	0.3657	0.2087	0.0632	0.0207		0.0

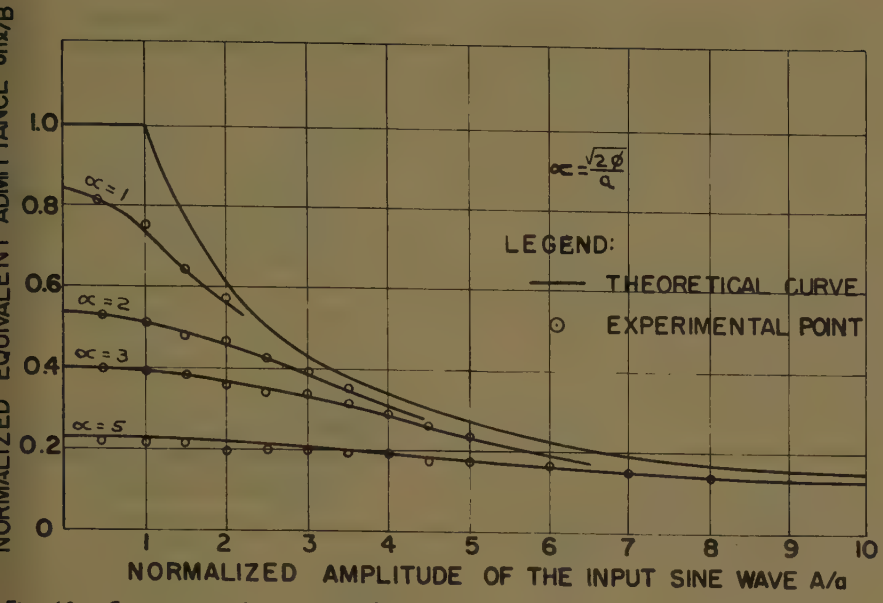


Fig. 13. Comparison of experimental and theoretical equivalent admittances of limiter

ith no noise at the input to the non-nearity, the amplitude and frequency f the system's output hunt, if any, may e determined by using the well-known escribing-function technique. The in-ersections of the polar plot of $G(j\omega)$ are ivestigated as is the negative of the escribing-function's reciprocal on the omplex plane of the nonlinearity. As ould be remembered, the describing unction is a special case of the equivalent dmittance.

What will happen if Gaussian noise ith a mean-square value of ϕ_0 is injected the input to the nonlinearity? This oise is going to affect the characteristic f the nonlinearity and, hence, the ampli-ude of the output hunt. Using the same ype of argument which justified replac-ng the nonlinearity with its describing unction when the input was a pure sine ave, we can now justify replacing the onlinearity with its equivalent admit-nce, since the input is a pure sine wave ogether with Gaussian noise.

In the latter case, the amplitude and equency of hunt is determined by in-estigating the intersection of the polar lot of the linear part—system $G(j\omega)$—nd the reciprocal's negative of the quivalent admittance. As the equiv-ent admittance of every nonlinearity onsidered in this paper is real and posi-ve, the reciprocal's negative of the quivalent admittance will lie along the egative real axis in the complex plane. he frequency of auto-oscillations is there-re determined by the intersection of $(j\omega)$ with this axis. Since this point does ot vary with changes in A or ϕ_0, the equency of auto-oscillations is not fected by the injection of noise. For a rticular mean-square value of noise, the amplitude of auto-oscillations may be determined from the intersection point.

From the standpoint of signal stabiliza-tion, however, our sole interest in the mean-square value of noise lies at the nonlinearity input, where hunting can be reduced to a desired value. Suppose the original hunt of the system is d_1 units, and this is to be reduced to d_2 units. A hunt amplitude of d_2 units at the system output will result in an amplitude of A_2 units at the input to the nonlinearity. This value is easily calculated when the actual trans-fer functions of the linear elements in the loop are known. The value of ϕ_{01} now sought for is one which causes $J_{n1}(A_1,\varphi_0)$ to intersect $G(j\omega)$.

The procedure to follow is quite sim-ple, as will be made clear after considering an example, based on the nonlinear feed-back system which is shown in Fig. 16. Using the describing function for the limiter, it can be shown that the ampli-tude and frequency of hunt of the system output are 63.5 units and 10 radians per second, respectively. As equivalent-ad-mittance curves for the limiter are shown in Fig. 13, the output hunt is seen to be 58 units and 32 units for $\sqrt{2}\phi_0/a=3$ and 5, respectively. Assuming that the noise fed back is negligible, then the mean-square value of noise at the limiter input is the same as at the system input.

Verification of Equivalent Admittance

An analog computer was used for experimentally verifying the equivalent admittance of the limiter. This verifica-tion should also serve to verify the equivalent gain, since this gain was used extensively in deriving equivalent-admit-tance curves.

The output of a commercial white-noise generator was passed through a second-order shaping filter and was added to the output of a sine-wave generator. The output of the adder was thus the sum of a sine wave and essentially Gaussian noise, which sum was the input to a limiter circuit. The fundamental component of the limiter output was measured with a sharply tuned filter. From this, the equivalent admittance of the limiter was evaluated. The experimentally deter-mined points appear as dots in Fig. 13, which shows close agreement between theoretically determined curves and these points.

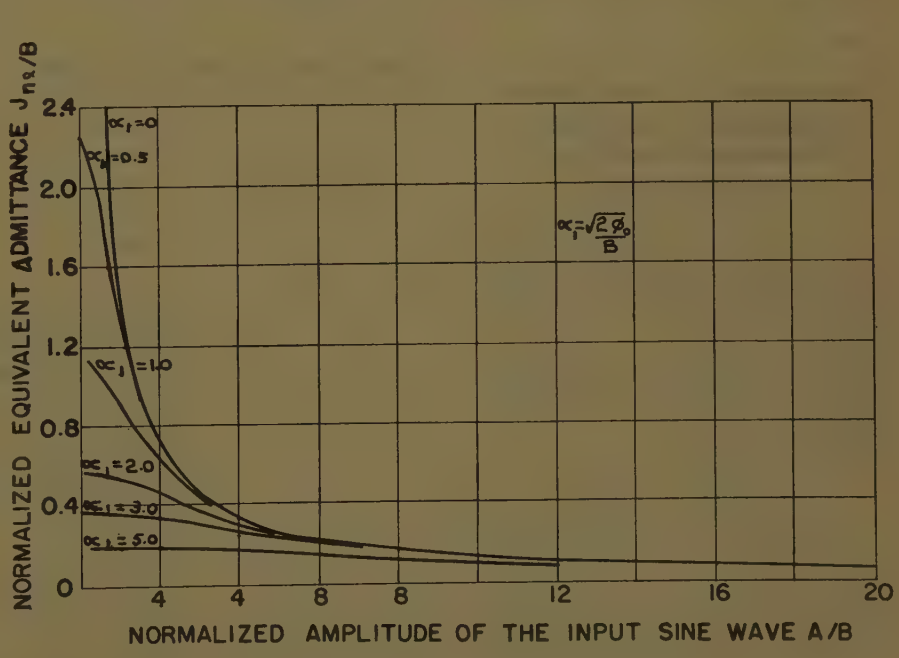


Fig. 14. Equivalent admittance $J_{n1}(A, \phi_0)$ of ideal relay

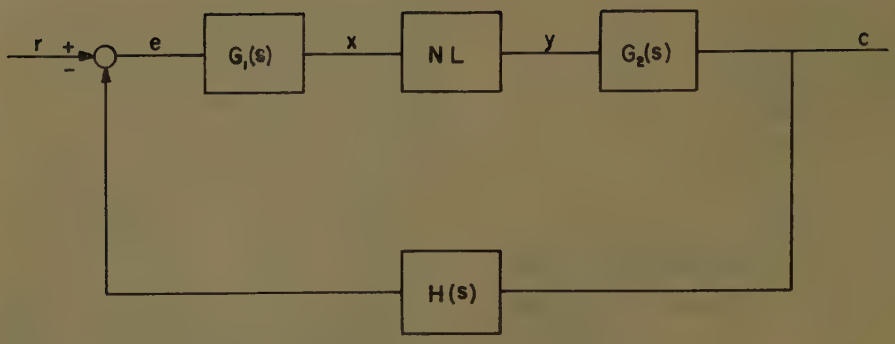


Fig. 15. General type of nonlinear feedback system

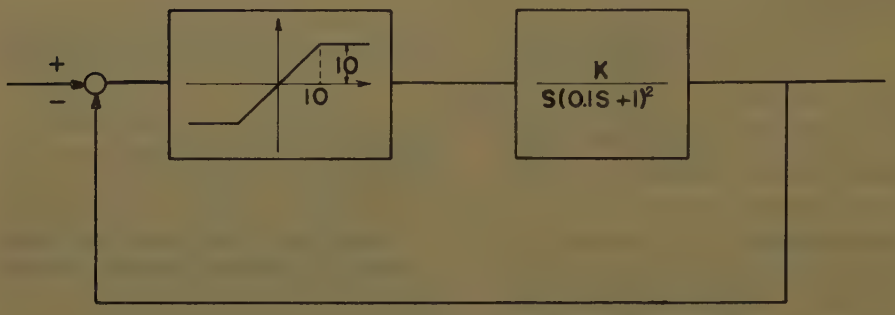


Fig. 16. A particular nonlinear feedback system

Amplitude of Auto-Oscillations of a Closed-Loop System

The closed-loop system in Fig. 16 was simulated on an analog computer. Noise with different mean-square values was injected at the system input for various values of the gain constant K, and amplitudes of the output hunt were measured. Results of this experiment, together with the theoretically calculated points, are shown in Fig. 17. Close agreement is again evident between the theory and experiment.

Appendix I. Input-Output Characteristics of Certain Piecewise-Linear-Type Nonlinearities

Input-output characteristics of the piecewise-linear-type nonlinearities shown in Figs. 1 through 7 are here described as a function. The alphabetical classification of each nonlinearity is also included.

1. The input-output characteristic of type A nonlinearity, illustrated in Fig. 1, is described by equation 14

$$y=f(x)=\begin{cases} m_2(x-a)+B & x\geq a \\ m_1(x-b) & b<x<a \\ 0 & -b\leq x\leq b \\ m_1(x+b) & -a<x<-b \\ m_2(x+a)-B & x\leq -a \end{cases} \quad (14)$$

2. The input-output characteristic of type B nonlinearity (limiter with dead band), seen in Fig. 2, is described by the following equation:

$$y=f(x)=\begin{cases} B & x\geq a \\ m_1(x-b) & b<x<a \\ 0 & -b\leq x\leq b \\ m_1(x+b) & -a<x<-b \\ -B & x\leq -a \end{cases} \quad (15)$$

3. The input-output characteristic of type C nonlinearity (limiter), Fig. 3, is

$$y=f(x)=\begin{cases} B & x\geq a \\ m_1x & |x|<a \\ -B & x\leq -a \end{cases} \quad (16)$$

4. Type D nonlinearity, shown in Fig. 4, has the input-output characteristic

$$y=f(x)=\begin{cases} m_2(x-a)+B & x\geq a \\ m_1x & |x|<a \\ m_2(x+a)-B & x\leq -a \end{cases} \quad (17)$$

5. Fig. 5 shows the input-output characteristic of type E nonlinearity described by

$$y=f(x)=\begin{cases} m_2(x-a) & x\geq a \\ 0 & |x|<a \\ m_2(x+a) & x\leq -a \end{cases} \quad (18)$$

6. The input-output characteristic of type F nonlinearity (relay with dead band), Fig. 6, is

$$y=f(x)=\begin{cases} +B & x\geq a \\ 0 & |x|<a \\ -B & x\leq -a \end{cases} \quad (19)$$

7. The ideal relay or type G nonlinearity has the input-output characteristic shown in Fig. 7 and expressed by

$$y=f(x)=\begin{cases} +B & x>0 \\ 0 & x=0 \\ -B & x<0 \end{cases} \quad (20)$$

Appendix II. Expression for Equivalent Gains of Some Piecewise-Linear-Type Nonlinearities

Expressions for equivalent gains of type B through G nonlinearities are derived by considering them as modifications of type A. By properly modifying the right sides of equations 4 and 5, expressions are found for equivalent gains of required nonlinearities each of which is considered separately.

1. Type B nonlinearity (limiter with dead band) is reduced from type A when $m_2=0$ in A. If $m_2=0$ in the right sides of equations 4 and 5, the result is

$$g_b(m)=\frac{m_1}{m}\sqrt{\frac{\phi_0}{2\pi}}\left[\exp\left\{-\frac{(a+m)^2}{2\phi_0}\right\}-\exp\left\{-\frac{(a-m)^2}{2\phi_0}\right\}-\exp\left\{-\frac{(b+m)^2}{2\phi_0}\right\}+\exp\left\{-\frac{(b-m)^2}{2\phi_0}\right\}\right]+\frac{m_1}{2}\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)+\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right]+\frac{m_1}{2m}\left[a\left\{\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right\}-b\left\{\operatorname{erf}\left(\frac{b+m}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{b-m}{\sqrt{2\phi_0}}\right)\right\}\right]-\frac{m_1}{2}\left[\operatorname{erf}\left(\frac{b+m}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{b-m}{2\phi_0}\right)\right] \quad (22)$$

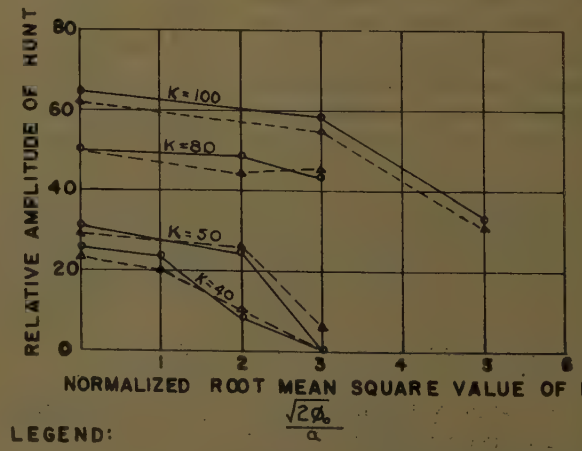


Fig. 17. Amplitude, oscillation of system output, Fig. 16

$$)=m_1\left[\operatorname{erf}\left(\frac{a}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{b}{\sqrt{2\phi_0}}\right)\right] \tag{22}$$

Type B nonlinearity reduces to the type imiter) when $b=0$ in the former. Let- $b\to 0$ in the right sides of equations 21 22 yields

$$:)=\frac{m_1}{m}\times\sqrt{\frac{\phi_0}{2\pi}}\left[\exp\left\{-\frac{(a+m)^2}{2\phi_0}\right\}-\exp\left\{-\frac{(a-m)^2}{2\phi_0}\right\}\right]+\frac{m_1}{2}\times\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)+\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right]+\frac{m_1}{m}\times\frac{a}{2}\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right] \tag{23}$$

$$)=m_1\operatorname{erf}\left(\frac{a}{\sqrt{2\phi_0}}\right) \tag{24}$$

Type A nonlinearity reduces to type D n $b=0$ in the former, yielding equation 25 n $b\to 0$ in equations 4 and 5

$$)=\frac{m_1-m_2}{m}\sqrt{\frac{\phi_0}{2\pi}}\left[\exp\left\{-\frac{(a+m)^2}{2\phi_0}\right\}-\exp\left\{-\frac{(a-m)^2}{2\phi_0}\right\}\right]+m_2+\frac{m_1-m_2}{2}\times\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)+\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right]+\frac{m_1-m_2}{m}\times\frac{a}{2}\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right] \tag{25}$$

equation 26

$$)=m_2\left[1-\operatorname{erf}\left(\frac{a}{\sqrt{2\phi_0}}\right)\right]+m_1\left[\operatorname{erf}\left(\frac{a}{\sqrt{2\phi_0}}\right)\right] \tag{26}$$

Type D nonlinearity reduces to the type hen $m_1=0$ in the former and when $m_1\to 0$ he right sides of equations 25 and 26

$$)=-\frac{m_2}{m}\sqrt{\frac{\phi_0}{2\pi}}\left[\exp\left\{-\frac{(a+m)^2}{2\phi_0}\right\}-\exp\left\{-\frac{(a-m)^2}{2\phi_0}\right\}\right]+m_2-\frac{m_2}{m}\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)+\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right]-\frac{m_2}{m}\times\frac{a}{2}\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right] \tag{27}$$

$$)=m_2\left[1-\operatorname{erf}\left(\frac{a}{\sqrt{2\phi_0}}\right)\right]$$

5. The expression for equivalent gain for type F nonlinearity (relay with dead band) can be derived more easily from the original definition than by considering it as a modification of any other type. Substituting the expression for $f(x)$ from equation 19 for A_1 in equation 3 and then substituting the resulting expression for A_v in equation 2 yields the following

$$g_f(m)=\frac{B}{2m}\left[\operatorname{erf}\left(\frac{a+m}{\sqrt{2\phi_0}}\right)-\operatorname{erf}\left(\frac{a-m}{\sqrt{2\phi_0}}\right)\right] \tag{28}$$

and

$$g_f(0)=B\sqrt{\frac{2}{\phi_0\pi}}\times\exp\left(-\frac{a^2}{2\phi_0}\right) \tag{29}$$

6. Now type F (ideal relay) reduces to the type G nonlinearity when $a=0$ in the former. Letting $a=0$ in the right side of equations 28 and 29 yields

$$g_g(m)=\frac{B}{m}\operatorname{erf}\left(\frac{m}{\sqrt{2\phi_0}}\right) \tag{30}$$

and

$$g_g(0)=B\sqrt{\frac{2}{\pi\phi_0}} \tag{31}$$

Appendix III. Average Output Value of Polynomial-Type Nonlinearity

In this derivation, rather than consider the general polynominal-type nonlinearity described by equation 6, the two special cases described by equations 7 and 8 are viewed separately.

For case 1, n odd, substitutions from equations 1 and 7 in equation 3 will produce the following

$$A_v=\frac{c}{\sqrt{2\pi\phi_0}}\int_{-\infty}^{\infty}x^n\exp\left\{-\frac{(x-m)^2}{2\phi_0}\right\}dx \tag{32}$$

By suitable change of variables and manipulation of the results, it can be shown that equation 32 yields

$$A_v=\frac{cm^n}{\sqrt{\pi}}\sum_{k=0,2,4}^{n-1}\frac{n!}{(n-k)!}\frac{1}{k!}\left(\frac{\sqrt{2\phi_0}}{m}\right)^k\gamma\times\left(\frac{k+1}{2}\right) \tag{33}$$

where Γ represents the gamma function.

For case 2, n even, substitutions from equations 1 and 8 in equations 3 will produce

$$A_v=\frac{-c}{\sqrt{2\pi\phi_0}}\int_{-\infty}^{0}x^n\exp\left\{-\frac{(x-m)^2}{2\phi_0}\right\}dx+\frac{c}{\sqrt{2\pi\phi_0}}\int_{0}^{\infty}x^n\times\exp\left\{-\frac{(x-m)^2}{2\phi_0}\right\}dx \tag{34}$$

By suitable changes of variables and manipulation of the results, equation 34 yields

$$A_v=\frac{c}{\sqrt{\pi}}m^n\left[\sum_{k=0,2,4}^{n}\frac{n!}{(n-k)!}\frac{1}{k!}\left(\frac{\sqrt{2\phi_0}^k}{m}\right)\times 2\int_0^{\frac{m}{\sqrt{2\phi_0}}}z^k\exp(-z^2)dz+\sum_{k=1,3,5}^{n}\frac{n!}{(n-k)!}\frac{1}{k!}\left(\frac{\sqrt{2\phi_0}}{m}\right)^k\times 2\int_{\frac{m}{\sqrt{2\phi_0}}}^{\infty}z\exp(-z^2)dz\right] \tag{35}$$

References

1. Signal Stabilization of a Control System, **R. Oldenburger.** *Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 79, no. 8, Nov. 1957, pp. 1869–72.

2. Signal Stabilization of a Control System, **R. Oldenburger, C. C. Liu.** *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 78, May 1959, pp. 96–100.

3. On Asynchronous Action, **N. Minorsky.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 259, no. 3, Mar. 1955, pp. 209–19.

4. Signal Stabilization of Self-Oscillating Systems, **R. Oldenburger, T. Nakada.** *Transactions*, Professional Group on Automatic Control, Institute of Radio Engineers, New York, N. Y., vol. AC-6, no. 3, Sept. 1961.

5. A General Method of Deriving the Describing Function of a Certain Class of Non-linearities, **R. Sridhar.** *Ibid.*, vol. AC-5, no. 2, June 1960, pp. 135–41.

6. On Some Nonlinear Phenomena in Regulatory Systems (in Russian), **L. C. Goldfarb.** *Automatika i Telemekanika*, Moscow, USSR, vol. 8, Sept.-Oct. 1947, pp. 349–83; also in Frequency Response (book, in English), **R. Oldenburger,** editor. The Macmillan Company, New York, N. Y., 1956, pp. 239–59.

Discussion

R. L. Cosgriff (Ohio State University, Columbus, Ohio): The authors are to be congratulated for focusing attention on characteristics of nonlinear systems excited by random signals. This discussion will review similar work and deal with extensions of the method presented in this paper.

A classic in the control field, which tells how to linearize the characteristics of a nonlinear device by a secondary or dither signal, is reference 1, and a general linearizing process is described by Loeb in reference 2.

The problem of random signals discussed in this paper is so closely associated with the gain function defined by Booton[3] that their small signal gain $g_1(0)$ is identical to Booton's gain function. Much of the paper's development parallels that of reference 4, although the techniques differ for obtaining the deterministic portion of the output signal.

Both treatments produce identical results for nonlinearities of the form $y=x^n$ when n is odd. Reference 4 also gives the series form for the terms when n is even, and allows determination of the output's expected value for nonlinear blocks, characterized by more general types of nonlinearity

than were possible with the author's development. For example, the expected value of y for

$$y=x(dx/dt)^2$$

and for

$$x=m(t)+n(t)$$

where $m(t)$ is a known signal and $n(t)$ is random with mean zero is determined by first substituting for x, giving

$$y=[m(t)+n(t)]\ [(dm/dt)^2+ 2\frac{dm}{dt}\frac{dn}{dt}+(dn/dt)^2]$$

The expected values of y become

$$y_e=m(t)\left[\sigma_z^2+\left(\frac{dm(t)}{dt}\right)^2\right]$$

Here σ_z^2 is the variance of dn/dt. For $m(t)=A\cos\omega t$, the value of y's fundamental component is

$$\left(\sigma_z^2+\frac{\omega^2}{4}\right)\cos\omega t$$

The authors' development of the gain function for certain nonlinear elements, for sinusoidal input signals, and for a random signal—based upon the concepts of Kochenberger and others—should prove to be useful; see Figs. 13 and 14. I regret that curves for relay with dead band were not included. However, the omission is understandable when considering the difficulty of presenting the data alone.

References

1. Fundamental Theory of Servomechanisms (book), L. A. McColl. D. Van Nostrand Company, Inc., Princeton, N. J., 1945, pp. 79–87.

2. A General Linearizing Process for Nonlinear Control Systems, J. M. Loeb. Automatic and Manual Control (book). Academic Press, Inc., New York, N. Y., 1952, pp. 275–83.

3. Analysis of Nonlinear Control Systems with Random Inputs, R. C. Booton. *Proceedings*, Nonlinear Circuit Analysis Symposium, April 24, 1953, pp. 369–92.

4. Nonlinear Control Systems (book), R. L. Cosgriff. McGraw-Hill Book Company, Inc., New York, N. Y., 1958, pp. 270–80

R. Oldenburger and R. Sridhar: Prof. Cosgriff's discussion adds greatly to the value of our paper. Certain developments cited by him parallel part of our work, although our main interest centers around stabi ing a feedback control system in limit cy operation by injecting a random sig This specific problem is not satisfacto explained in any of the references mentio in the discussion.

The theory of signal stabilization can extended rather easily to nonlinearit more general than the ones considere This is evidenced by Professor Cosgr example of the nonlinearity whose out involves the input and its derivative. H ever, this is a matter of details rather t generalities. The theory's extension to s tems with nonsingle valued and other co plicated nonlinearities warrants furt work. A simple example is a system wh has an element with hysteresis.

The theory developed to explain sig stabilization with random inputs eviden may be extended to yield more informat It can, for example, obtain stability infor tion for nonlinear systems with a n pickup in the loop. (See reference 1)

Reference

1. Stability of a Nonlinear Feedback Sys in the Presence of Gaussian Noise, R. Srid R. Oldenburger. *Paper no. 61-JAC-5*, Amer Society of Mechanical Engineers, New York, N

A Graphical Method for Finding the Frequency Response of Nonlinear Closed-Loop Systems

A. S. McALLISTER
STUDENT MEMBER AIEE

SEVERAL methods[1–7] have been developed for finding the steady-state sinusoidal response of closed-loop systems containing nonlinear elements. Most of these methods require either unnecessarily severe assumptions or long iterative procedures. This paper presents a graphical procedure for finding the frequency response of a large class of systems for which the describing function provides an adequate representation for the nonlinear elements. Attention is restricted to frequency-insensitive zero-memory nonlinear elements. The method is direct, requiring only an understanding of the point of view, and it can be extended to cover systems containing more than one nonlinear element.

A Determination of Closed-Loop Sinusoidal Response

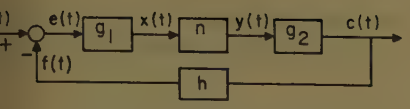
Consider the single-loop nonlinear system shown in the block diagram of Fig. 1. The blocks marked g_1, g_2, and h are linear elements whose input–output relations can be expressed by the usual transfer functions using the Laplace variables, $G_1(s)$, $G_2(s)$, and $H(s)$. The block marked n is a nonlinear element whose instantaneous output can be expressed as a single-valued odd function of its instantaneous input, that is, this block is a frequency-independent zero-memory nonlinear element. Under these conditions, if the system input $r(t)$ is a steady sinusoid, the other system signals e, x, y, c, and f are generally periodic but not necessarily sinusoidal. In an important class of control systems the linear elements have low-pass characteristics, so that even though y may be nonsinusoidal, its higher harmonics are filtered as the signal passes around the loop, making x very nearly sinusoidal. Under these conditions only the fundamental component of y is of importance in analyzing the closed-loop sinusoidal behavior. Thus, with the limitations of frequency insensitivity and zero memory, the nonlinear element can be approximated as a linear gain element whose gain $N(X)$ function of only the amplitude of the put X. This gain is the describing fu tion of the nonlinear element. Throu out this analysis it is assumed that input of the nonlinear element is sinuso and the describing function is an adequ representation of the nonlinearity. A ysis of the system can be carried if all the system signals are considere be sinusoids of the same frequency w the input r is sinusoidal.

It should be noted that even if the tem loop has a sufficiently low-pass c acteristic, so that the describing func method may be used for sinusoidal a ysis, the system output c need not ne sarily be a pure sinusoid but can con harmonics. If the values of any monic components of the output are sired, they can be approximately de mined by first assuming that the desc ing function gives the correct value This value is used to find the harmo in y. The harmonics in c can the found by applying $G_2(j\omega)$ with the ap priate values of ω.

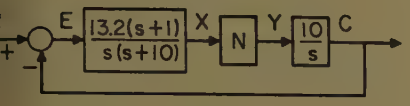
Now, if all the signals of the system sinusoidal, then the system is linear any given value of the gain, N, and

Paper 61-710, recommended by the AIEE Feed Control Systems Committee and approved b AIEE Technical Operations Department presentation at the AIEE-AIChE-ASME-IRE Joint Automatic Control Conference, Bou Colo., June 28–30, 1961. Manuscript subm September 8, 1960; made available for pri May 9, 1961.

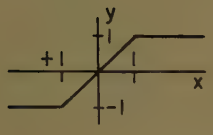
A. S. McAllister is an Associate Profess San Jose State College, San Jose, Calif.



1. Feedback control system containing a nonlinear element



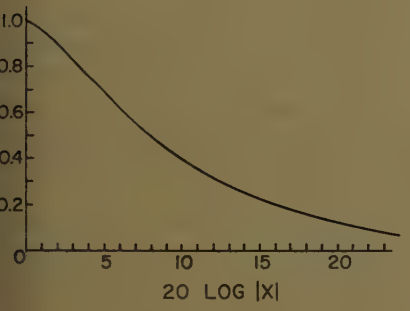
(A)



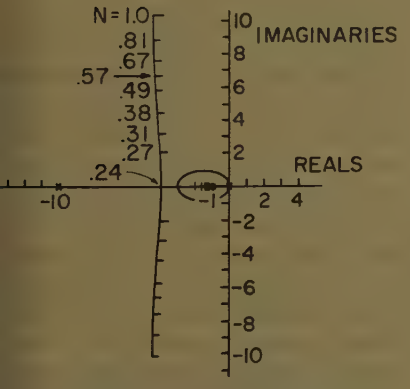
(B)

2. Specific example of nonlinear system

A—Block diagram
B—Idealized saturation



3. Describing function for idealized saturation of Fig. 2



ig. 4. Root locus for system of Fig. 2

plitude ratio and relative phase of C R can be obtained in the usual nner.

$$\frac{X}{R}(j\omega,N)=\frac{G_1(j\omega)N(X)G_2(j\omega)}{1+G_1(j\omega)N(X)G_2(j\omega)H(j\omega)} \quad (1)$$

is, the ratio C/R can be determined for given N and frequency ω. Un-

Fig. 5. Closed-loop frequency response of system of Fig. 2 for various values of N

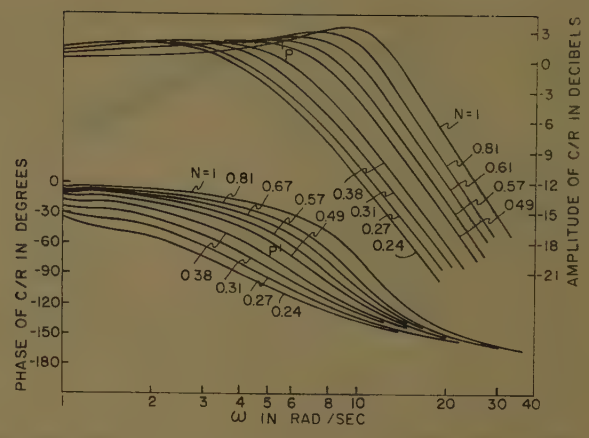
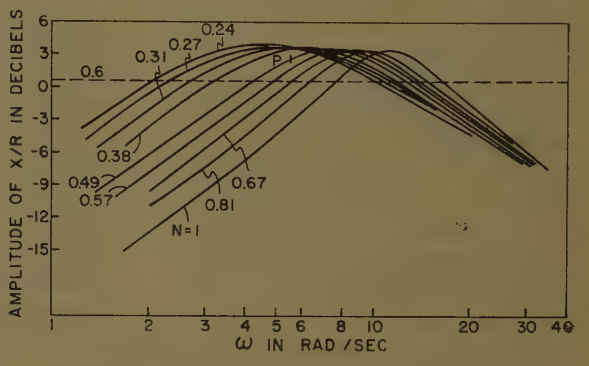


Fig. 6. X/R curves for system of Fig. 2



fortunately this is not the information that is desired; what is needed is the value of this ratio as a function of the amplitude and frequency of the input R. But the value of X (and therefore the value of N) is not immediately obtainable when R is given. Indeed, X is a very complicated function of R. The transfer function relating the two,

$$\frac{X}{R}=\frac{G_1(j\omega)}{1+G_1(j\omega)N(X)G_2(j\omega)H(j\omega)} \quad (2)$$

is a function of X. The solution of equation 2 for X, in terms of R, is usually prohibitively difficult even for the simplest nonlinearities.

However, if the point of view is slightly changed, a method which readily lends itself to graphical solution is presented. The solution of equation 2 is extremely difficult or sometimes impossible, when R is treated as the independent variable and the corresponding values of X and N are sought; however, the solution is simplified if X is taken as the independent variable. For a given amplitude of X, N is a constant and X/R is an ordinary linear closed-loop transfer function which can be solved as a function of frequency ω in any of the usual ways.[1,8,9] Thus, R can be found as a function of frequency for the given X. If this is done for a sufficient number of values of X, the difficulty will have been circumvented. R will be known for every frequency, ω, and every gain value N. Now, when the amplitude and frequency of an input, R, are given, the accumulated data are examined and the corresponding value of N is found. This value in equation 1 gives the desired output C.

It is desirable to have a complete set of data relating C to R for all amplitudes and frequencies of R; this may be accomplished by utilizing all the data that relate N to R in the repeated solutions of equation 1. At first glance this appears to be a tremendous task, but with systematic procedures and graphical techniques it is no more difficult than the plotting of the closed-loop frequency response of a linear system for several values of loop gain. The following section illustrates such a graphical method of solution.

Example Using a Graphical Approach

A specific example of the system of Fig. 1 is shown in Fig. 2(A) where n is taken as idealized saturation as shown in Fig. 2(B). The describing function, N, for saturation has been given by many authors[10] and is shown in Fig. 3. Though much of the work of this section is carried out in terms of the complex Laplace variables, it must be remembered that under the assumption of the describing function the analysis is only valid when x is less than the saturation value 1 or when s is purely imaginary, $s=j\omega$, that is, when the signals are all steady-state sinusoids of the same frequency.

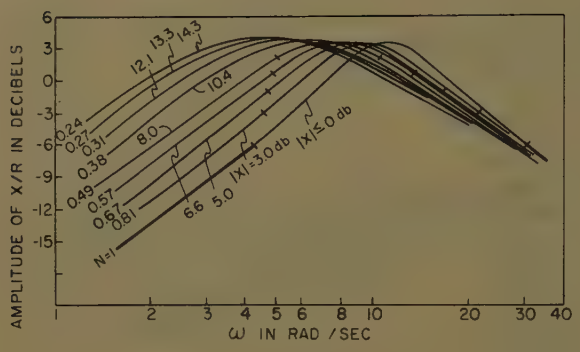


Fig. 7. Replot of Fig. 6 with $|R|=6$ db points identified

The two transfer functions which are of interest can now be written

$$\frac{C}{R}=\frac{\dfrac{13.2(s+1)}{s(s+10)}N\dfrac{10}{s}}{1+\dfrac{13.2(s+1)}{s(s+10)}N\dfrac{10}{s}}=\frac{N132(s+1)}{s^2(s+10)+N132(s-1)} \quad (3)$$

$$\frac{X}{R}=\frac{\dfrac{13.2(s+1)}{s(s+10)}}{1+\dfrac{13.2(s+1)}{s(s+10)}N\dfrac{10}{s}}=\frac{132s(s+1)}{s^2(s+10)+N132(s+1)}=\frac{C}{R}\frac{s}{N10} \quad (4)$$

To plot curves of these functions for $s=j\omega$, the roots of the denominator common to both functions must be determined. The location of the roots can be obtained as a function of N by means of the root locus[1,11]. The root locus for the system under consideration is shown in Fig. 4. The linear gain 132, is selected so that the complex roots have a damping factor, $\zeta=0.4$, when operating in the linear region. The values of the describing function, N, are marked on the complex roots in Fig. 4. The real roots are not marked. The data from Fig. 4 are given in Table I, where

$$s^2(s+10)+132N(s+1)=(s+p_3)(s+2\zeta\omega_n s+\omega_n^2) \quad (5)$$

$132N$ is the loop gain, N the describing function, ζ the damping factor of the complex pair of roots, ω_n the undamped natural frequency of the complex roots, $-p_3$ the location of the third root, and $|X|$ the value of input to the nonlinearity corresponding to N as obtained from Fig. 3. It can be seen in Table I that the values of gain were chosen so that the values of the damping factor, ζ, would be simple. This was done so that standard templates could be used to draw the frequency response curves.

For constant values of N, all the poles and zeros of equations 3 and 4 are known, and frequency response curves can be drawn by plotting log-amplitude and phase versus log-frequency. Asymptotic approximations to these curves are not sufficient because these closed-loop functions involve complex poles and because fairly accurate frequencies at which certain amplitudes occur are wanted. However, if templates or similar devices are used, it is reasonably easy to add the effects of the various poles and zeros graphically to obtain the final curves. The curves used in this paper were drawn with the aid of a set of normalized curves which were placed under the final curve sheet and added as they were traced. The curves of equations 3 and 4 are shown in Figs. 5 and 6. Since only the amplitude of X is needed for determining the value of the describing function, the phase of X/R is not necessary and therefore has not been drawn.

In drawing Fig. 6 a simplification was used which might be worth mentioni From equation 4 it is seen that

$$\frac{X}{R}=\left(\frac{C}{R}\right)\left(\frac{s}{10N}\right)=\left(\frac{C}{R}\right)\left(\frac{j\omega}{10N}\right)$$

Therefore, for any given N and ω, value $|X/R|$ in Fig. 6 can be determi from the corresponding value of $|C/R$ Fig. 5 by a simple multiplication (by a ing logarithms). So Fig. 5 was dra first using the normalized curves and t Fig. 6 was obtained by adding 20 log $10N$ to the amplitude curves point point. This procedure is generally ea as the relation between X and C is an o loop function and usually is much simp than the closed-loop functions.

To find the amplitude of R associa with each point on these curves note identity

$$20\log|X/R|=20\log|X|-20\log|R|$$

so that the input $|R|$ for each curv Fig. 6 is given by

$$20\log|R|=20\log|X|-20\log|X/R|$$

Thus, in order to find the input for point on the amplitude curves of Fig take the value of $|X|$ in decibels (corresponding to the N of the curve question and subtract from it the va $|X/R|$ given by the curve itself at t point.

To illustrate this consider point P in 6 for which $\omega=5.5$ rad/sec (radians s second), $|X/R|=2.7$ db, and $N=0$ The value of $|X|$ corresponding to N is found to be $X=8.0^5$ db in Tabl therefore, from equation 7 the amplit of the input for this point is

$$|R|=8.0^5-2.7=5.3^5 \text{ db}$$

It follows that the points on the am tude and phase curves of Fig. 5 hav the same N and ω as point P must have the same input value and they h been so marked.

In the foregoing example: some a trary point on the $|X/R|$ curves picked; the values of ω, N, $|X/R|$, and for this point were found; and t values were used to determine $|R|$ for point. Actually, since contours of for constant R will eventually be wan it would be nicer if the required value could be selected and then the point the $|X/R|$ curves which corresponded t could be identified. These points c be transferred to the C/R curves as fore and then connected to give a con for the given input value of R. This be done. Since N is constant along one of the curves of Fig. 6, the valu $|X|$ is constant along each curve ex

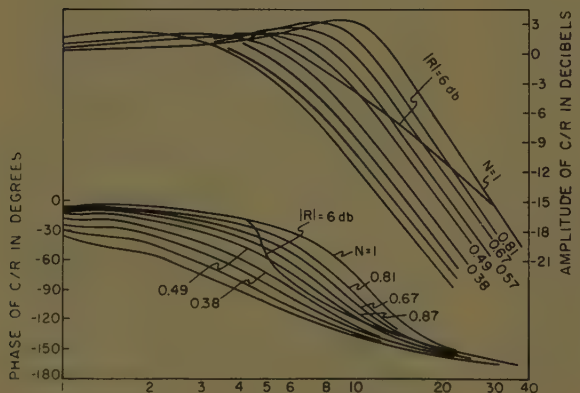
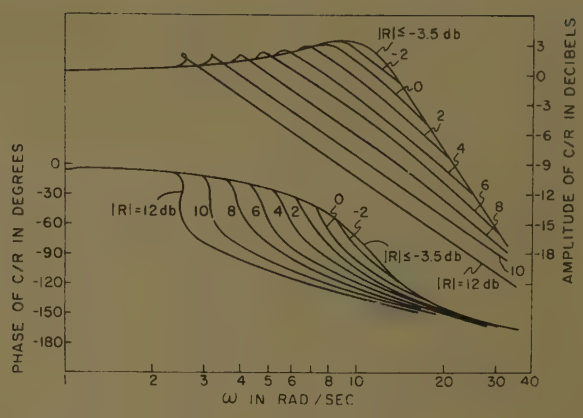


Fig. 8. Replot of Fig. 5 with $|R|=6$ db curve drawn

ble I. Root Locations for $s^2(s+10)+132N(s+1)=0$

2N	N	ζ	ω_n, sec^{-1}	P_s	|X|, db
2	1	0.4	11.2	1.08	≤0
75	0.815	0.45	9.7	1.11	3.05
0	0.675	0.5	8.9	1.13	5.0
5	0.57	0.555	8.05	1.16	6.65
4	0.49	0.6	7.3	1.195	8.05
0	0.38	0.7	6.2	1.29	10.4
15	0.315	0.8	5.4	1.43	12.1
55	0.27	0.9	4.65	1.63	13.35
2	0.24	1.0	4.0	2.00	14.35

Fig. 9. Closed-loop frequency response of system of Fig. 2 for various values of input



=1. Let us try to find all the points r which $|R|$ has the value, say 6 db. ow on the $N=0.57$ curve, for example, e see from Table I that $|X|=6.6^5$ nerefore, from Equation 6, $|R|$ will equal at the point on this curves for which

$|X/R|=6.6^5-6=0.6^5$ db

at is, at points where the $N=0.57$ rve in Fig. 6 crosses the 0.6^5-db line. nere are two such points; see Fig. 6. milar calculations with other values of show where other $|R|=6$ db points cur; these have been identified in Fig.

To prevent any confusion from Fig. 7, should be pointed out that, although e assumed restrictions on the type of nlinearity being studied force $N(X)$ be a single-valued function of the nplitude of X, the converse is not nerally true. Thus, in the present case, =1 whenever $|X|\leq 1$ and the linear =1 curve is valid whenever $|X|\leq 0$ db. , from equation 6, $|X/R|$ is given by the =1 curve at all points for which $|X/R|\leq |R|$. This means that a single point the linear curves may be valid for more an one input amplitude. This should use no more difficulty than when the stem is entirely linear if the nature the nonlinearity is kept in mind roughout the analysis.

The points on Fig. 7 can be transferred the curves of Fig. 5, as before, which s been done in Fig. 8, and have been nnected to form the contours of con- nt $|R|=6$ db. With reference to these rves it can be seen that the input is w enough at low frequencies so that the tput is able to follow fairly well. The or, E, is so small that X does not satu- e. However, as the frequency gets gher with R constant, the system lags re and more increasing the error until finally becomes greater than 1, then urates and N decreases. In fact, as n in Fig. 8, N decreases rather rapidly h frequency in this range. This might qualitatively explained by the fact t as X starts to saturate the output has even more trouble following the input than it would have if the system were linear and the amplitude and phase of the output cannot change much if Y is able to change only slightly. Therefore, the error increases very rapidly causing N to decrease rapidly over a short frequency range. (In fact, as will be seen later, N can change instantaneously in some instances.)

Eventually, the low-pass characteristic of G_1 takes over and X slowly starts to decrease again. In this region, therefore, Y is very nearly constant and the output is determined by G_2 (in this case, a straight line with slope of about −6 db per octave). Finally, G_1 makes X so small that it no longer saturates and operation is again linear.

The curves for other values of input can be obtained in the same way. Fig. 9 shows several such curves drawn for values of $|R|$ chosen 2 db apart. This system has been simulated on an analog computer and the results are compared with Fig. 9 in Appendix I.

The reader may be interested in the little curlicues that occur in Fig. 9 for high input values. The system was not particularly chosen to show the phenomenon and actually the curve for $|R|=12$ db is a little unusual in that it shows a closed figure, but this is the so-called jump resonance. If the input amplitude were held at this high value and the frequency increased slowly from a low value, the system would operate at first as previously explained. But as operation continues along the curve, the system reaches a point where the frequency cannot keep increasing with operation staying smoothly on the curve. The output must then jump to the next portion of the curve, with X increasing and N decreasing discontinuously, where things again run smoothly. On the other hand, if the frequency were decreased from a high value, and the input amplitude were decreased from a high value and the input amplitude were kept constant at 12 db, a jump would again occur, but at a different, lower frequency. Thus, for some frequencies there are two possible output amplitudes. (Actually the curve shows three values at these frequencies, but as Levinson[6] points out, the third is unstable.) This phenomenon is discussed by Lozier,[2] Prince,[5] Levinson,[6] and others. Unfortunately, the approximation used by both Lozier and Prince leads to the conclusion that a system, such as the one discussed here, would always exhibit jump resonance for all input amplitudes which is not the case.

Fig. 9 provides all the necessary information for analyzing the sinusoidal behavior of the linearized system of Fig. 2. The thoroughness of the data is limited only by the amount of work the engineer is willing to do. It hardly seems necessary to draw curves closer than 2 db as the depicted curves give a good picture of how the system behaves. Furthermore, in any practical system, it would probably be expected that the input would never exceed some given value, in which case it would be unnecessary to draw curves for higher input values. However, curves can be obtained for as high an input amplitude as desired by drawing enough N curves.

Finally, in some applications it might be desirable to have normalized output $|C|$ plotted rather than the output-input ratio. This can be accomplished by simply adding the appropriate value of $|R|$ to each of the curves of Fig. 9, which has been done in Fig. 10. If the phase position of the input is taken as zero, then the phase of the output is the same as that of C/R shown in Fig. 9. Since the output of the saturation nonlinearity, y, approaches a square wave of amplitude 1 as the input, x, gets large, the curves of Fig. 10 approach a limit which has been included on the graph.

Systems Containing More Than One Nonlinear Element

Although considerable work has been and is being done toward finding the

stability, transient response, sinusoidal response, and response to statistically described inputs of systems, which are adequately characterized as having only one significant nonlinear element, little has been attempted on systems in which more than one of the nonlinearities are important.[7] However, such systems do occur, such as a servomechanism in which an amplifier, whose input is limited, drives a motor whose nonlinear characteristics cannot be neglected. This section describes how the previously discussed method can be adapted for the sinusoidal analysis of systems of this sort.

It is not hard to see why much study has not been done on systems containing several nonlinearities. In the first place, quite often all but one of the nonlinearities are negligible, even though this is not always the best design. But more important, it is hard enough to work with only one nonlinear element and adding a second not only doubles the work but compounds the difficulties.

This paper utilizes the describing function to characterize the nonlinear elements. The input to each nonlinearity is assumed to be a sinusoid when the system input is a sine wave. With the single nonlinearity it is fairly common for the open-loop characteristic to be low-pass because of mechanical devices in the loop, slow amplifiers, etc., but this is not sufficient when there are several nonlinear elements. If describing functions are used, there must be a low-pass filter between every two nonlinearities so that the harmonics created by one are sufficiently filtered to make the input to the next essentially sinusoidal. Thus, a describing function method generally will be less valid in the multiple nonlinearity case and the accuracy may leave something to be desired. However, if there is at least one integration, $(1/s)$, between the nonlinearities, the method should provide at least an indication of the sinusoidal response characteristics of the system. (Actually, although the system given in the following example was picked completely at random, the results of an analog computer study are quite close to the analytical calculation.)

A simple system containing two nonlinear elements is shown in Fig. 11 where the linear elements are expressed in terms of their Laplace transfer functions, G_1, G_2, G_3, and H, and the nonlinear elements are given by their describing functions N_1, N_2. Assuming that the describing functions are valid, it is seen that

$$\frac{C}{R}=\frac{G_1N_1G_2N_2G_3}{1+G_1N_1G_2N_2G_3H} \tag{8}$$

$$\frac{X_1}{R}=\frac{G_1}{1+G_1N_1G_2N_2G_3H} \tag{9}$$

With only one nonlinearity the poles and zeros of the closed-loop functions were found as a function of the gain N and curves of C/R and X/R were plotted. Since N and X had a known functional relationship (the describing function), it was possible to find the value of R corresponding to each point on the X/R curves. These points could then be identified on the C/R curves which provide the final curves of C/R for constant R. In the present case, the poles and zeros may still be found as a function of gain, but the gain is now the product of the two describing functions, N_1N_2, and is not simply related to any of the signal amplitudes.

Therefore, to find the input value for each point, it is necessary to know how N_1N_2 is related to the amplitude of X_1. This may be done in two ways. The more obvious way is to choose several values of X_1 first. For any given amplitude of X_1, the amplitudes of N_1 and Y_1 are independent of the frequency and are obtainable from the describing function N_1 (X_1). For any given frequency of Y the amplitude ratio, $|X_2/Y_1|$, is ind pendent of amplitude and is known fro the transfer function G_2. Thus, for a given amplitude and frequency of X_1, it possible to find the corresponding amp tude of X_2 and therefore N_2. For t given X_1 the product, N_1N_2, is the known. If this process is carried throug for a large number of values of X_1, curv could be drawn from which X_1 could determined by any given N_1N_2 and ω.

This relationship may also be obtaine by choosing X_2 first. For any given X the value of N_2 is independent of fr quency and can be determined fro the describing function $N_2(X_2)$. Thu for any given amplitude and frequency X_2, it is possible to find the correspondi amplitude of Y_1, by using G_2 inverse Now, though it is not usually thought in this way, the describing function mathematically just as much a function an element's output as its input. Ce tainly, under the assumption that t input is a sinusoid, if the fundamental the output is given, the value of the d scribing function can be found. Ther fore, for the given amplitude and fr quency of X_2, if Y_1 is known, then bo N_1 and X_1 can be determined and t corresponding value of the product, N_1 can be calculated. Again, if this is do for several values of X_2 and ω, curv could be drawn relating X_1, N_1N_2, a ω. This second approach is used he because it yields data in a form that is little easier to handle; thus, more a curacy is possible with less effort.

There appears to be no obvious reas why X_1/R curves had to be used to fi the necessary relations for the det mination of the desired input-outp curves. With a single nonlinear eleme it was very convenient to do so since was constant along any one of the However, in the present case this is true. From the previous discussion it fairly obvious that it would not be a more difficult to relate N_1N_2, ω, and then to relate N_1N_2, ω, and X_1. So, Y_2 curves could be used instead of the X_1

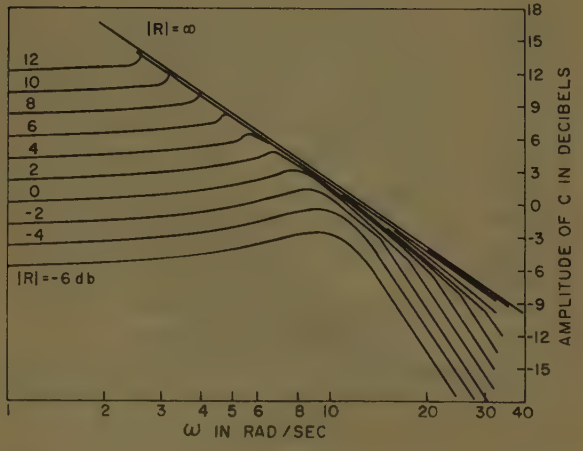


Fig. 10. Output of system of Fig. 2 for various values of input

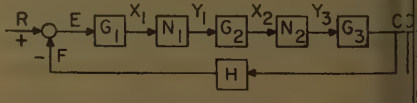


Fig. 11. System containing two nonlin elements

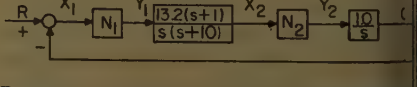


Fig. 12. Specific example of system conta ing two nonlinear elements

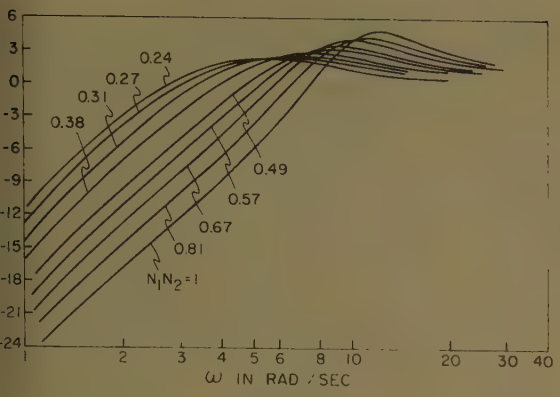


Fig. 13. X_1/R curves for system of Fig. 12

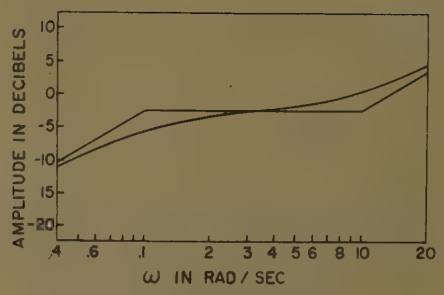


Fig. 14. Amplitude of G_2^{-1}

$$= \frac{j\omega(j\omega+20)}{13.2(j\omega+1)}$$

rves, and a little more effort would re-al the relation between N_1N_2, ω, and itself, in which case the original C/R rves could be used without any re-urse to a second set such as those of uation 9. Actually the amount of ork needed in any of these cases is just out the same; the choice of method based on greater convenience and other ch considerations. In the following ample, X_1 is used rather than Y_2 or C it is somewhat easier and more ac-rate.

To illustrate how the input–output rves are obtained, consider the system shown in Fig. 12 where both N_1 and N_2 are taken to be saturation as shown in Fig. 2(B). The gain values have been so chosen that when operating in the linear region on N_1, ($X_1 \leq 1$), the system is the same as that studied in the previous section. Therefore, the curves of Fig. 5 are the C/R curves for this system if N is replaced by N_1N_2. The curves of the amplitude of X_1/R are shown in Fig. 13.

Now, since X_2 will be chosen and then Y_1 determined, a plot of

$$\left|\frac{Y_1}{X_2}\right| = \left|\frac{1}{G_2(j\omega)}\right| = |G_2^{-1}(j\omega)| = \left|j\omega\frac{(j\omega+10)}{13.2(j\omega+1)}\right| \quad (10)$$

will be needed. Because N_1 and N_2 are only amplitude sensitive, just the amplitude of G_2^{-1} is required and this is drawn in Fig. 14. Further, a curve relating N_1 to Y_1 is needed; this is just a redrawing of Fig. 3 using the fact that

$$Y_1 = N_1(X_1)X_1 \quad (11)$$

and is shown in Fig. 15.

Now the values of $|X_2|$ must be selected. The values of the product, N_1N_2, that are wanted are the values of N given in Table I. Since $N_1=1$ for $|X_1| \leq 1$, let values of $|X_2|$ be chosen such that N_2 takes on the values of N given in Table I. Then N_1N_2 will have these values when N_1 is not saturating. Also, select a few more values between these just to increase the number of points on the curves. Table II provides the values of $|X_2|$ and N_2 that will be used in this discussion.

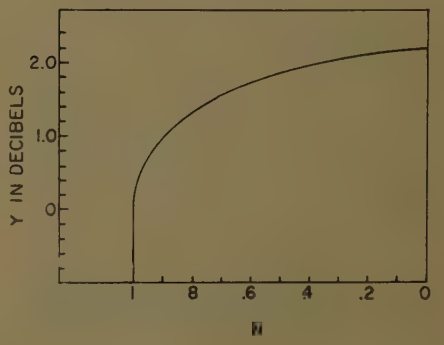


Fig. 15. Inverse describing function for saturation

To obtain curves of Y_1 as a function of frequency for the values of N_2 given in Table II, the values of $|X_2|$ are added to the $|G_2^{-1}|$ curves of Fig. 14. The resulting curves are shown in Fig. 16.

Before proceeding, it will be instructive to discuss Fig. 16 briefly. It is helpful to be aware of some of the limits of operation that exist. First, it is seen that $N_2=1$ for all points below the $|X_2|=0$ db curve and that $N_2 \leq 1$ at all points above. Thus, N_2 is limiting only at those points above the $|X_2|=0$ db curves. Also, $N_1=1$ for all points below the $|Y_1|=0$ db curves; the first nonlinearity limits only above this line. Finally, no matter how large X_1 is, the largest Y_1 can be is the fundamental of a square wave of magnitude 1, therefore, the maximum value of $|Y_1|$ is 2.10 db. Consequently, the portions of the curves above the $|Y_1|=2.10$ db line are impossible.

ble II. Describing Function for Saturation from Fig. 3

$\|X_2\|$, db	N_2
0	1
1.8[5]	0.90
3.0[5]	0.81[5]
3.5[5]	0.78
3.9[5]	0.75
4.5[5]	0.71
5.0	0.67[5]
5.8[5]	0.62
6.6[5]	0.57
7.3[5]	0.53
8.0[5]	0.49
9.1	0.44
10.4	0.38
12.1	0.31[5]
13.3[5]	0.27
14.3[5]	0.24

ble III. Points on Fig. 16 Where $N_1N_2=0.49$

N_2	N_1	$\|Y_1\|$, db	ω, rad/sec	$\|X_1\|$, db
....	0.49	1.86.....	13.1	8.0
90	0.54[5]....	1.78.....	8.7	7.1
81[5]....	0.60	1.72.....	5.68	6.2
78	0.63	1.68.....	4.36	5.7
75	0.65	1.64.....	3.42	5.3
71	0.69	1.58.....	2.52	4.8
67[5]....	0.72[5]....	1.51.....	1.81	4.3
62	0.79	1.35.....	1.23	3.4
57	0.86	1.14.....	0.95[5]....	2.4[5]
53	0.92[5]....	0.83.....	0.90	1.5[5]
49	1	≤ 0	≤ 0.61	≤ 0
44				
38				
31[5]	impossible points for $N_1N_2=0.49$;			
27				
24	$N_2<N_1N_2$, since $N_1 \leq 1$.			

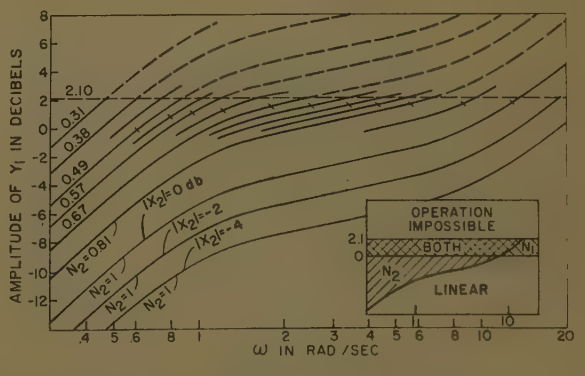


Fig. 16. Amplitude of Y_1 for system of Fig. 12. Insert shows where the two nonlinear elements saturate

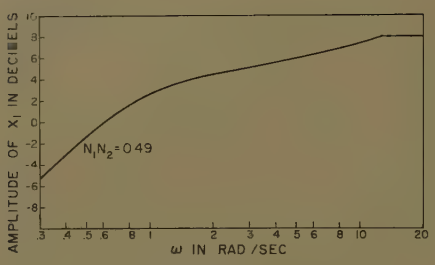


Fig. 17. Amplitude of X_1 for system of Fig. 12 and $N_1N_2=0.49$

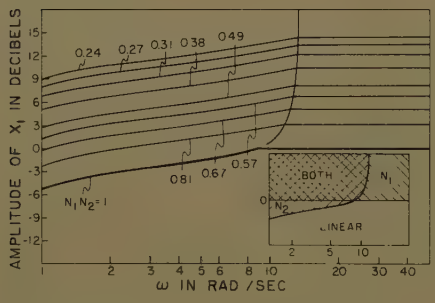


Fig. 18. Amplitude of X_1 for system of Fig. 12. Insert shows where the two nonlinear elements saturate

Thus, the regions of Figure 16 in which each nonlinear element is saturating may be determined.

Now the points in Fig. 16 where the product N_1N_2 has the values used in Figs. 5 and 13 may be found. In the region where only N_2 saturates, $N_1=1$, $N_1N_2=N_2$, and the desired N_1N_2 curves are those already drawn in Fig. 16. In the region where only N_1 saturates, the product, N_1N_2, is entirely determined by N_1 which in turn is a function of $|Y_1|$. In this region the constant N_1N_2 curves are constant $|Y_1|$ curves which are horizontal lines.

In the region where both nonlinearities saturate, some point-by-point calculation will be needed to find the desired points. For instance, let us find the points where $N_1N_2=0.49$ by using Table III. The first column gives the values of N_2 on the various curves of Fig. 16 as given in Table II. The second column is merely $N_1N_2=0.49$ divided by N_2. The third column is obtained from the inverse describing function of Fig. 15 and gives the value that $|Y_1|$ must have to yield the N_1 given in the previous column. The first three columns contain all the necessary data to find the $N_1N_2=0.49$ points on Fig. 16. N_1N_2 must equal 0.49 at each point where the N_2 curve in column 1 intercepts the value of $|Y_1|$ given in column 3. These points have been identified in Fig. 16.

Now the curves relating X_1, ω, and N_1N_2 are needed, so columns 4 and 5 have been added to Table III. The frequencies in the fourth column are read directly from Fig. 16 and are the frequencies at which the various points occur. The fifth column is obtained from the describing function of Fig. 3. Now the curve of $|X_1|$ versus ω may be drawn for $N_1N_2=0.49$ and is given in Fig. 17. For $\omega \leq 0.61$ rad/sec this curve is identical with the $N_2=0.49$ curve of Fig. 16 because $X_1=Y_1$. For $0.61 \leq \omega \leq 13.1$ rad/sec the curve is plotted from columns 4 and 5 of Table III, and for $\omega \geq 13.1$ rad/sec $N_1=0.49$ and $|X_1|=8.0$ db.

This curve is repeated in Fig. 18 together with similar curves obtained in the same manner for the other values of N_1N_2. Fig. 18 is drawn to the same scale as the $|X/R|$ curves of Fig. 13. It is important for subsequent development that these two curve sheets be drawn to the same scale.

Now it is possible to find the value of the input R at any point on the curves of Fig. 13. These curves give $|X_1/R|$ versus ω for a certain set of values of gain N_1N_2, while Fig. 18 gives curves of $|X_1|$ versus ω for the same values of N_1N_2. Therefore, from the identity

$$20 \log |R| = 20 \log |X_1| - 20 \log |X_1/R| \tag{12}$$

the value of the input, $|R|$, for any given N_1N_2 and ω is the value of $|X_1|$ in db, given by Fig. 15, minus the value of $|X_1/R|$ in db, given by Fig. 13. There are several ways to perform this subtraction. It is certainly valid just to pick points at random and use equation 12 to find $|R|$ at these points, but as with a single nonlinearity, it is better and possible to choose the value of input first and then find the appropriate points on the curves. The first thing that comes to mind is some sort of trial-and-error method where a few points are selected with an eye toward converging on the desired point by the application of equation 12, but a more systematic method is available if the curves of $|X_1/R|$ and $|X_1|$ are drawn on tracing paper using the same scale. By use of equation 12, for example, the 6 db points occur where the $|X_1/R|$ curve and the $|X_1|$ curve for the same N_1N_2 are 6 db apart, with the latter greater. So, one set of curves is placed over the other with the 0-db lines coincident and a pair of dividers is set at 6 db, these points can be found with very little difficulty.

Actually there is an even better way of doing this same thing. If the $|X_1/R|$ curves are placed over the $|X_1|$ curves so that the 0-db line of the $|X_1/R|$ curve coincides with the 6-db line of the $|X_1|$ curves, then equation 12 is automatically satisfied at intersections of similar N_1N_2 curves for $|R|=6$ db. An attempt to show this process is given in Fig. 19 where the resulting 6-db points are shown. These points may then be transferred to the C/C curves and connected to form the $|R|=6$-db contour, which has been done in Fig. 20.

Similar statements may be made about the $R=6$-db curves of Fig. 20 as were made about Fig. 8. If the amplitude of the input is kept constant at $R=6$ db and the frequency is slowly increased from a low value, then at the low frequencies, the error is small and the system operates linearly along the $N_1N_2=1$ curve. Eventually, the error increases so that one of the nonlinearities begins to saturate. In this particular instance, the second nonlinearity, N_2, saturates first. Now, with only N_2 saturating, this system is the same as that with only one nonlinearity, and in this portion of the curve is identical with that of Fig. 8. As before, the error increases and the gain decreases rapidly because Y_2 cannot increase rapidly. Now, the first nonlinear element, N_1, begins to saturate, but it does so in such a way that the system becomes unstable immediately. The error must increase instantaneously and the gain decrease instantaneously with a slight increase of frequency so that the output jumps to the upper section of the curve. This is the same sort of effect as the jump resonance mentioned previously but the effect is caused here by the presence of two saturations.

With further increase of frequency, the

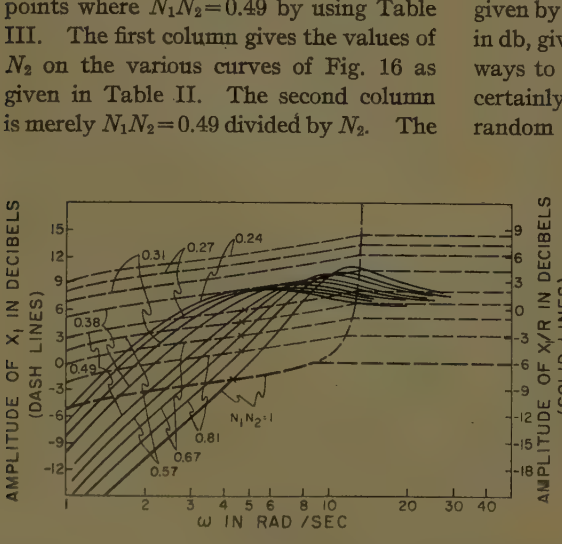


Fig. 19. Determination of $|R|=6$ db points

w-pass characteristics of both the $G_2 = .2\ (s+1)/s(s+10)$ and $G_3 = 10/s$ begin take effect and both X_2 and C decrease gardless of the fact that the error, X_1, mains large. Eventually, X_2 becomes small that it no longer saturates and ly X_1 remains saturated. The point at iich this occurs lies on the curve sep- ating the regions where both N_1 and N_2 turate and only N_1 saturates. This rve was taken directly from Fig. 18 d is shown in Fig. 20. At higher fre- iencies, the output C continues to de- ease and X_1 approaches the input R. ius, the first nonlinearity, N_1, remains turated at all high frequencies for $|R| =$ db.

In a similar manner, curves like those of g. 20 can be obtained for other input lues. Fig. 21 shows several of these rves. This system has been simulated an analog computer and the results e compared in Fig. 21 in the Appendix.

Since this section has dealt with graph- al solution it is well to discuss briefly the curacy of the results. At several points the work, pairs of curves have been aphically added to obtain new curves. ius, the accuracy obtained by such chniques is questionable and scepticism not without ground; however, the real iestion is how much accuracy is war- nted. The method is based upon the sumption that the inputs to the various onlinear elements are all sinusoidal when e input is sinusoidal. This is only proximately true even in the case of a ngle nonlinearity and when there are veral nonlinear elements, the approxi- ation is much worse. In most instances, e method will, at best, probably give ly an indication of how the output aries with the amplitude and frequency the input. The curves should be awn as accurately as possible under the sumptions, but it hardly seems neces- ry to require more accuracy than can obtained from graphical techniques.

onclusions

A method has been developed for find- g the response of certain types of non- iear feedback control systems to steady- ate sinusoidal input signals. The ethod makes use of the describing func- on to represent the nonlinear elements. miliar linear frequency response tech- ques are used to draw families of curves om which the input-output data of e nonlinear system are deduced. For y given system the method is no more fficult than the plotting of the frequency sponse curves of the same system with e nonlinearities replaced by linear gain elements for several values of gain. The nonlinear response data are obtained in the form of contours of fundamental (and any desired harmonics; see Appendix II) of the output signal as a function of frequency and the amplitude of the system input signal. Complete response data are available for any desired frequency or input amplitude from a single set of curve sheets. A new set of curves is not needed for every different input amplitude.

The method is illustrated by means of two examples which are worked out in full. As a check on the accuracy of the results the two systems have been simulated on an analog computer. The data from the computer roughly show the method to be as accurate as the computer itself. The second example involved a system containing two nonlinear elements and, although the method would not be expected to give as accurate results for this case as for the single nonlinearity case, the analog results agree very closely with the curves obtained analytically.

Appendix I. Analog Computer Simulation

The systems used as examples have been simulated on an analog computer. The results of the simulation are shown in Figs. 22 and 23.

Appendix II. Harmonics in the Output

As was stated previously, although the input to the nonlinear element may be so nearly sinusoidal that the describing function may be used, the system output may contain significant harmonics and it may be desirable to calculate their values. This is not at all difficult to do when the describing function is indeed applicable. Truxall[1] has shown that for idealized saturation if the input is a sinusoid of amplitude $|X|$ the nth harmonic of the output is

$$|Y^{(n)}| = \frac{2|X|}{\pi n}\left\{\frac{\sin(n-1)t_2}{n-1} + \frac{\sin(n-1)t_2}{n-1}\right\} \qquad (13)$$

for $|X|>1$ and $n=1,3,5\ldots$, where $t_2 = \text{arc sin } |1/X|$ and the notation $Y^{(n)}$ is taken to mean the nth harmonic of the output y. Thus the harmonics of y are known and since the transfer function G_2 from y to the system output c is linear the harmonics in the output can easily be calculated.

To illustrate how this is done let us consider the third harmonic output of the system considered in the preceding section.

$$\left|\frac{Y'''}{X}\right| = \frac{1}{3\pi}\left(\sin 2t_2 + \frac{1}{2}\sin 4t_2\right) \qquad (14)$$

Now Table IV can be filled in. Column 1 contains the values of N used throughout this analysis; column 2 gives the corresponding values of $|X|$ from Table I; column 3 provides the corresponding values of $|Y'''/X|$ from equation 13; and column 4 contains the sum of columns 2 and 3. For any given input frequency, ω, therefore, the

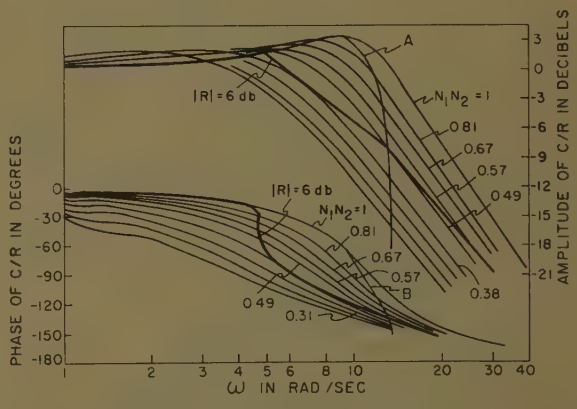


Fig. 20. Fig. 5 replotted with $|R| = 6$ db curve drawn. Curves A and B show where N_2 just begins to saturate

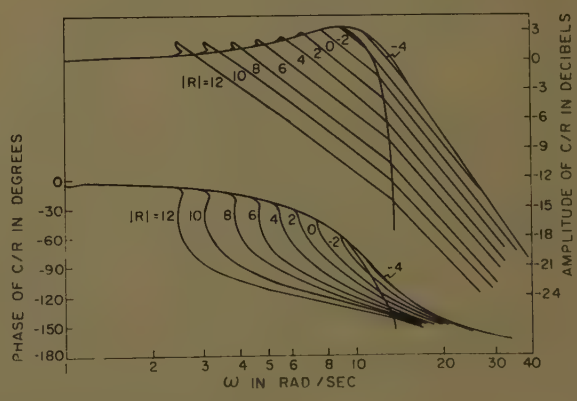


Fig. 21. Closed-loop frequency response of system of Fig. 12 for various values of input

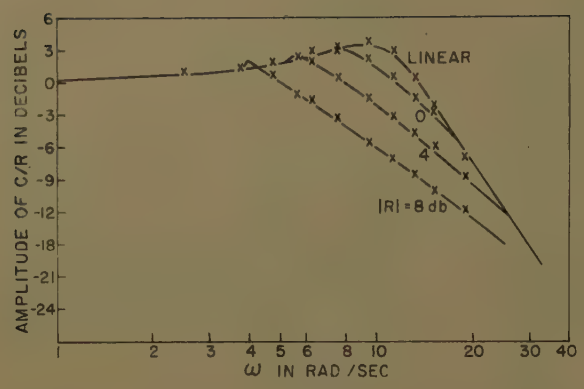


Fig. 22. Results of analog simulation of system of Fig. 2

Solid lines repeated from Fig. 9
x represents analog data

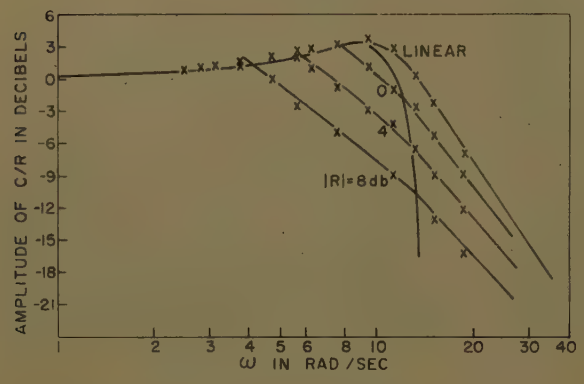


Fig. 23. Results of analog simulation of system of Fig. 12

Solid lines repeated from Fig. 21
x represents analog data

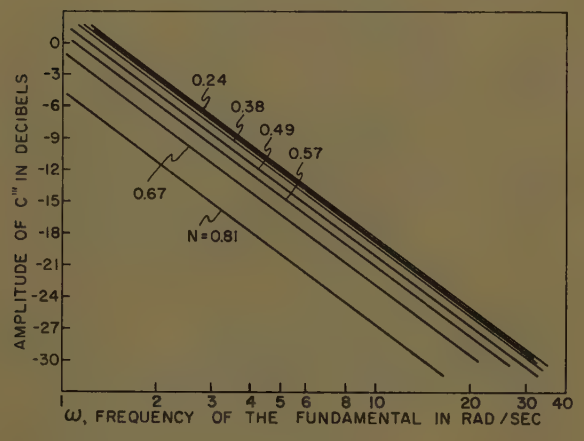


Fig. 24. Third harmonic output of system of Fig. 2 for various values of gain N

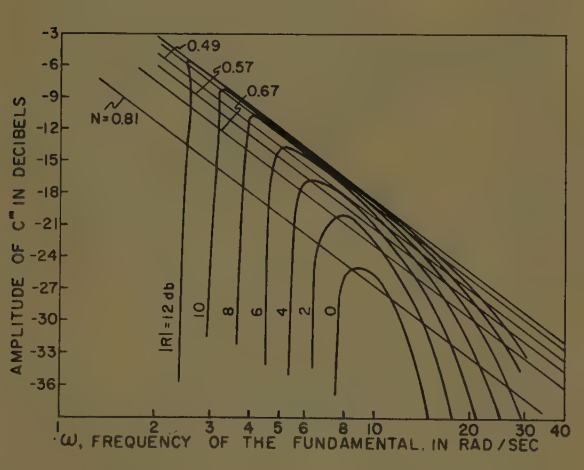


Fig. 25. Third harmonic output for various values of system input

third harmonic component of system output (see Fig. 2) is

$$C''' = G_2(j3\omega)Y''' = \frac{10}{j3\omega}\,Y''' \qquad (15)$$

The curve of $|G_2(j3\omega)| = |10/j3\omega|$ is a straight line of slope minus 20 db, per decade. Therefore, for constant values of $|Y'''|$ corresponding to constant values of N, equation 15 is a series of straight lines as shown in Fig. 24 where the abscissa is the frequency of the system input, that is, the frequency of the fundamental of the output.

Table IV. Third Harmonic in Saturation

N	|X|, db	$\left\|\frac{Y'''}{X}\right\|$, db	|Y|, db
1	≤0	$-\infty$	$-\infty$
0.81[5]	3.0[5]	−19.2	−16.
0.67[5]	5.0	−17.3[5]	−12.
0.57	6.6[5]	−17.3[5]	−10.
0.40	8.0[5]	−17.7	−9.
0.38	10.4	−19.0[5]	−8.
0.31[5]	12.1	−20.4	−8.
0.27	13.3[5]	−21.5	−8.
0.24	14.3[5]	−22.4	−8.

Now, from the work in the paper, the valu of N and ω, corresponding to given values input R, are known and can be identified Fig. 24. This has been done on Fig. 25 a the points have been connected to fo contours of constant input.

The relative phase of third harmo requires discussion. From equation 15 it seen that C''' lags Y''' by $\pi/2$ radia This means that when the third harmonic y is going through zero with positive slo the third harmonic of c is going through negative maximum. To see how this lates to the phase of the fundamental n that the output waveshape of the nonlin element has sine symmetry (is an odd fu tion) about the point where it pas through zero. Therefore, when the fun mental of y is going through zero with po tive slope, all the harmonics also p through zero and with positive slo Furthermore, at this same instant fundamental of the output, c, is go through its negative maximum as are all t harmonics in the output. Thus, the outp waveshape has cosine symmetry (is an ev function) about this point and the outp wave is going through its negative ma mum. This means that the output wa shape will always be peaked, that is, m pointed than a sine wave.

Comparing Fig. 25 with Fig. 10 it is se that the amplitude of the third harmonic the output is more than 19 db lower th that of the fundamental. That is, the th harmonic output has about 1/10 t amplitude of the fundamental or small Furthermore, the higher harmonics wo be even less significant. The output wa form is itself not very different from bein sinusoid, which was assumed.

In addition, the factor $1/s$ in G_1 wo decrease the harmonics still further so t the input to the nonlinearity is indeed v nearly sinusoidal. This harmonic analys as the analog study of Appendix I confir shows that neglecting harmonics is p missible and the method presented sho give highly accurate results.

References

1. Automatic Feedback Control System S thesis (book), J. G. Truxall. McGraw-Hill B Company, Inc., New York, N. Y., 1955.

2. A Steady State Approach to the Theory Saturable Servo Systems, *Proceedings*, In tute of Radio Engineers, New York, N. Y., PGAC-1, May 1956, pp. 19–39.

3. An Analytic Method for Finding Closed-Loop Frequency Response of N linear Feedback-Control Systems, K. Og

EE Transactions, pt. III (*Power Apparatus and stems*), vol. 76, Nov. 1957, pp. 277–85.

A General Method for Analyzing and nthesizing the Closed Loop Response of a near and a Nonlinear Servomechanism, H. El-Sabbagh. *WESCON Convention Record*, stitute of Radio Engineers, pt. 4 (*Automatic ntrol*), 1957, pp. 58–77.

Generalized Method for Determining the osed-Loop Frequency Response of Non-near Systems, L. T. Prince, Jr. *AIEE Transac-ns*, pt. II (*Applications and Industry*), vol. 73, pt. 1954, pp. 217–23.

Some Saturation Phenomena in Servo-chanisms with Emphasis on the Tachometer Stabilized System, E. Levinson. *Ibid.*, vol. 72, Mar. 1953, pp. 1–9.

7. Frequency Response of Nonlinear Closed-Loop Feedback Control Systems, S. L. Mikhail, G. H. Fett. *Ibid.*, vol. 77, Nov. 1958, pp. 436–38.

8. Principles of Servomechanisms (book), G. S. Brown, D. P. Campbell. John Wiley & Sons, Inc., New York, N. Y., 1948.

9. Servomechanism Analysis (book), G. J. Thaler, R. G. Brown. McGraw-Hill Book Company, Inc., New York, N. Y., 1953.

10. Sinusoidal Analysis of Feedback-Control Systems Containing Nonlinear Elements, E. Calvin Johnson. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 71, Jan. 1952, pp. 169–81.

11. An Extension of the Root Locus Method to Obtain Closed-Loop Frequency Response of Feedback Control Systems, A. S. Jackson. *Ibid.*, vol. 73, Sept. 1954, pp. 176–79.

12. Prediction of Transient Response of Nonlinear Servomechanism by Sinusoidal Analysis, H. C. Brearley, Jr. *Ph.D. Thesis*, The University of Illinois, Urbana, Ill., 1954.

13. A Frequency Response Method for Analyzing and Synthesizing Contactor Servomechanisms, Ralph J. Kochenburger. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 270–84.

Discussion

. Sridhar (Purdue University, Lafayette, id.): It seems quite obvious that most uthors who use the describing function of a onlinearity to obtain the closed loop fre-uency response of a nonlinear system do ot seem to appreciate the full limitations of e describing function. A sufficient condi-o.i for the describing function method of otaining the closed loop frequency response a nonlinear system to be valid is that the ystem be totally stable. Otherwise, it is ossible to obtain quite erroneous results sing this method. It is possible for a sys-m which is perfectly stable for no inputs, indicated by the describing function ethod, to break into auto-oscillations hen subjected to sinusoidal inputs, the equency of the auto-oscillations being in-ependent of the forcing frequency. Again, is possible to quench the auto-oscillations certain autonomous nonlinear systems by ubjecting them to sinusoidal inputs so that ie system will respond to the forcing fre-uency. This latter is, of course, the well-nown phenomenon of signal stabilization. It is high time that authors stopped pro-osing new methods of obtaining the frequency response of nonlinear systems using the describing function method. It will possibly be worth while to establish criterion which will tell when the describing function method is valid and when it is not.

The author appears to have naively used the describing function to obtain the frequency response of a nonlinear system with two nonlinearities. Actually, the validity of the describing function method to investigate the stability of such systems, even when autonomous, has not been established for a general case. I shall appreciate any comments the author has regarding the assumptions on the linear elements when he analyzed the system.

A. S. McAllister: I want to thank Mr. Sridhar for his thoughtful discussion. The points made by him are in most cases well taken. The blind use of linearization techniques can indeed lead to erroneous results. But this has long been the case whether using Hooke's law, studying Class *A* amplifiers or analyzing feedback control systems. The engineer should always keep in mind that he is making an approximation and should check his results and assumptions frequently and as carefully as possible. If a rigorous validity criterion that can be applied easily is available it would be well to use it.

All this certainly limits the general applicability of such techniques. However, it does not invalidate the usefulness of linearization. Indeed, linearization is usually easier and faster to use than the more elaborate and rigorous methods. There are times when the linearizing method will give an approximate answer where the exact method becomes so involved as to be utterly useless. The more rigorous methods simply have not yet been developed to the point where they can be used in the more complicated problems.

As for the assumptions made in analyzing a two-nonlinearity system, they are given in the third paragraph of the appropriate section of the paper and need not be repeated here. In this connection, it might be noted that Mikhail and Fett[1] also analyzed a two-nonlinearity system using the two describing functions. They obtained rather good results even though the system they studied had no linear elements, low-pass or otherwise, between the nonlinear elements.

Reference

1. See reference 13 of the paper.

On Stabilization of Feedback Systems Affected by Hysteresis Nonlinearities

A. K. MAHALANABIS
NONMEMBER AIEE

PRESENCE OF HYSTERESIS or hereditary types of nonlinearities ads to small-signal destabilization in rvomechanisms and other feedback con-ol systems and produces small-ampli-ide sustained oscillations or chattering, rticularly when precision requirements ake necessary the use of a high loop in. Some safeguards against this un-sirable characteristic are therefore pre-nted here.

Several methods of providing proper mpensation for hysteresis effects have en suggested.[1–5] The use of coulomb friction for this purpose has been investigated by this author,[6,7] and the present paper embodies results of further studies in this direction. Some nonlinear compensation schemes for improving the small-signal stability of feedback control systems are dealt with. They not only permit compensation for the action of hysteresis effects on stability but also result in improved response characteristics.

The general desirability of such nonlinear compensation methods in inherently nonlinear systems has been stressed, for example, by Truxal, who considered an interesting case of nonlinear compensation, around a feedback loop, for backlash effects.[8]

Hysteresis Effects in Simple Servo Systems

Simple electromechanical systems are considered here. Basic components of an electromechanical servomechanism, shown in Fig. 1, are: (1) the error-sensing unit, (2) the controller-amplifier unit, (3)

Paper 61-709, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE-AIChE-ASME-IRE-ISA Joint Automatic Control Conference, Boulder, Colo., June 28–30, 1961. Manuscript submitted September 6, 1960; made available for printing May 8, 1961.

A. K. Mahalanabis is with the University College of Technology, Calcutta, India.

The author is grateful to Professor J. N. Bhar and Dr. A. K. Choudhury for their assistance; also to the Ministry of Scientific Research and Cultural Affairs, Government of India, for providing funds during the progress of this work.

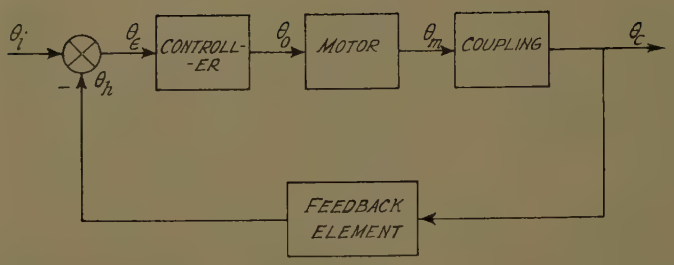


Fig. 1. Basic components of a servomechanism

the motor, (4) the output coupling unit.

One or more of these units generally have a hysteresis-type nonlinearity, whose source could be: (1) backlash in the output coupling unit, (2) the controller in a contactor system which incorporates a contactor, e.g., a biopolarized (electromagnetic) relay, possessing hysteresis, and (3) field-controlled electromagnetic devices, e.g., amplidyne generators in the controller unit or a field-controlled motor. That hysteresis may arise from sources 2 or 3 is evident. Backlash in the output coupling produces a hereditary effect when the load has negligible inertia and damping, equal to or greater than critical.

The functional characteristics arising from nonlinearity in different cases are illustrated in Figs. 2(A) through 2(D). Fig. 2(E) is an approximation of the magnetic hysteresis curves, which simplifies analysis by describing function techniques. Performance of the system shown in Fig. 1 is difficult to analyze when more than one form of hysteresis occurs simultaneously. Analyses of individual effects of these nonlinearities have, however, been made. The frequency-response approach, in which the nonlinearity is replaced by its describing function, has been found most convenient for stability analysis of nonlinear systems. The describing functions, Figs. 2(A), (B), and (E), have been discussed in published literature.[1,2,9] Although these functions differ in detailed properties, a common feature is the predominance of nonlinear effects at comparatively small-signal amplitudes, where a reduction in the magnitude of transfer gain is produced, together with a lagging phase shift. Both effects increase as the signal amplitude approaches the hysteresis width.

The phase lags cause the system's stability to be impaired at small-signal levels, producing sustained oscillations of essentially small amplitude in an otherwise stable system.

Stabilization of Oscillations Arising from Hysteresis Effects

Sustained oscillations outlined in the previous section can be prevented. This calls for reducing the loop-gain sufficiently to avoid an intersection in the gain-phase shift plane of the linear-frequency locus and the amplitude locus of the negative reciprocal of the relevant describing function. Sluggish response characteristics result from this procedure, however. The nonlinear-gain element suggested by Johnson[2] seeks to remedy this defect; see Fig. 3(A). But an extreme use of this plan, as exemplified by the dead-zone element shown in Fig. 3(B), may impair the static accuracy of the resulting system. Use of conventional phase-lead networks may also be somewhat unsatisfactory, as will be discussed subsequently.

Stabilization of oscillations in systems having hysteresis may be accomplished through improving system stability at small-signal levels without affecting the response characteristics at large-signal amplitudes. This is because hysteresis introduces a sort of variable damping with effective damping becoming small as the signal amplitude diminishes. A logical procedure for compensating for the

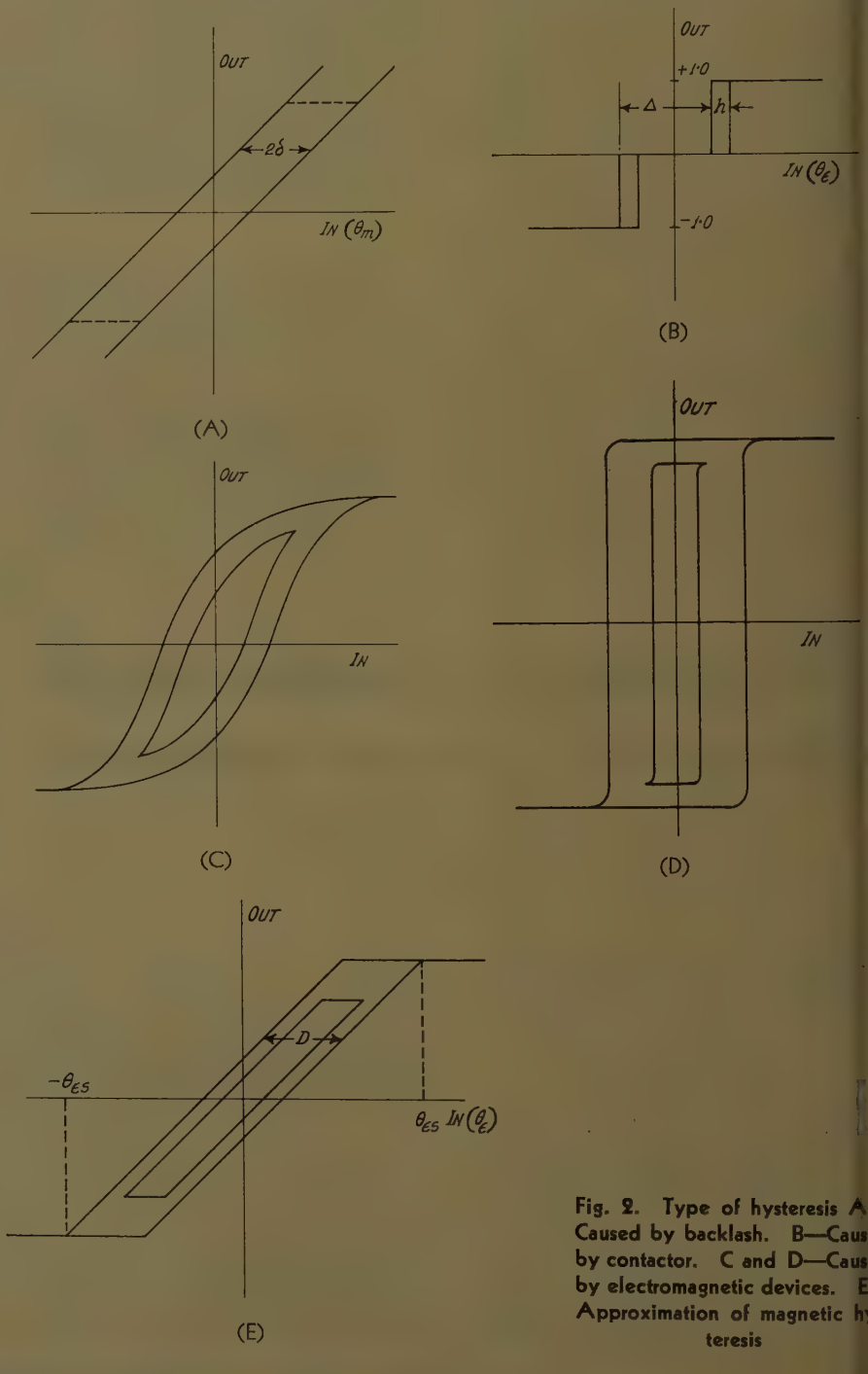
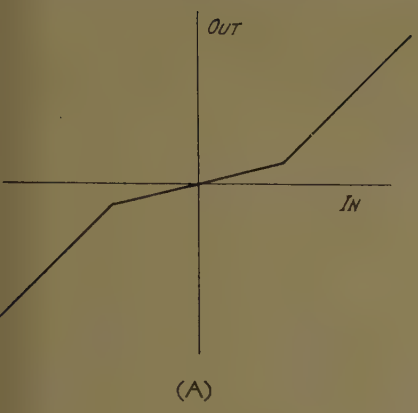
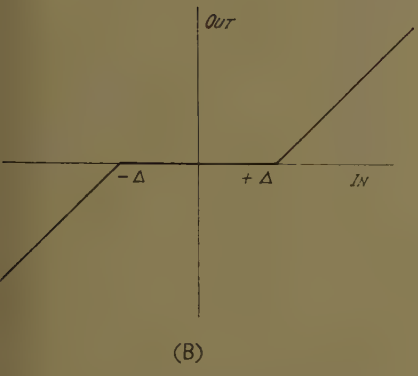


Fig. 2. Type of hysteresis A—Caused by backlash. B—Caused by contactor. C and D—Caused by electromagnetic devices. E—Approximation of magnetic hysteresis



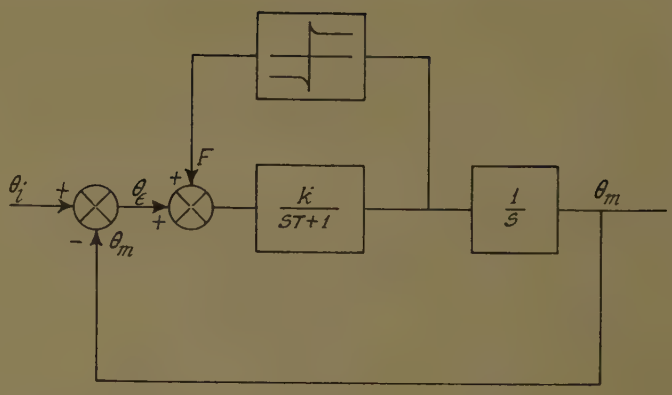


g. 3. A—Nonlinear-gain characteristics. B—Dead-zone characteristics

ysteresis effects is, therefore, to insert a omponent in the forward or feedback ath that, in absence of hysteresis, also roduces a variable damping, but with ne damping becoming larger as the signal nplitude falls.

The nonlinear-gain element, Fig. 3(A), certainly one possibility; others are scussed in subsequent sections. These tter fall into two groups: nonlinear rate edback and nonlinear phase-lead com-ensation. Feeding back a signal pro-ortional to some function of the system peed and introducing a cascade phase-ad network are widely practiced for im-roving a system's stability of response. ogically then, one or both of these tech-niques may be modified to suit the prob-lem under consideration.

Fig. 4. Schematic of system affected by coulomb friction



Nonlinear Rate Feedback

General Facts

Some recent investigations on the effects of coulomb friction in feedback systems having hysteresis nonlinearities reveal that this friction has beneficial effects on stability.[6,7] This is because it produces, in effect, a nonlinear feedback of the rate signal. See Fig. 4. However, presence of coulomb friction has a generally bad effect on static accuracy and causes a zone of steady-state errors. Stabilization of feedback systems containing hysteresis by intentional introduction of coulomb friction may thus be undesirable, particularly in high-precision systems.

The good effects of coulomb friction on small-signal stability of a system can, however, be realized in practice if the secondary feedback loop seen in Fig. 4 is not caused by the friction but represents a suitable external feedback of the rate signal. This is explained in discussing the case of a simple second-order servo having backlash hysteresis.

A Nonlinear Rate Feedback System

Arrangements of the proposed nonlinear rate feedback system is shown in Fig. 5(A). The secondary feedback loop, around the block representing the motor, results in the feedback of a nonlinear rate signal H, the form of which depends on the characteristics of the nonlinear block N. For the purpose of the present discussions, block N is supposed to have the characteristics shown in Fig. 5(B). This is realized in practice by placing a limiter ahead of a tachometer generator.

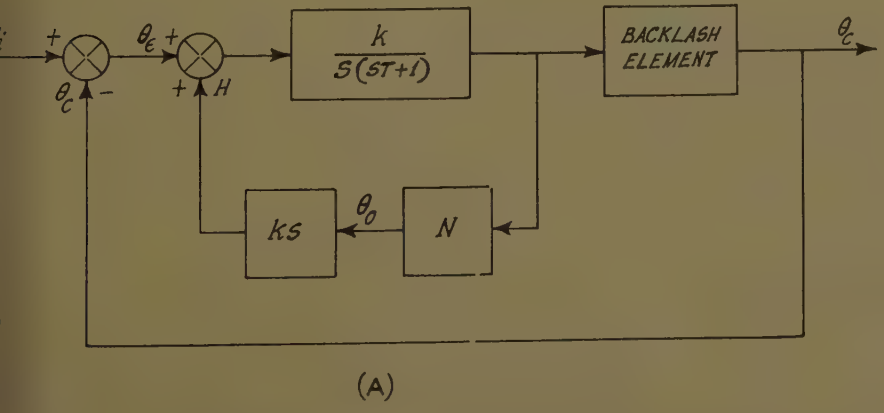
Analysis of Nonlinear Rate Feedback System

A frequency-response analysis of the system in Fig. 5(A) is possible by replacing nonlinearities with corresponding quasilinear describing functions. The equivalent system is shown in Fig. 6. The portion of Fig. 6 shown within the broken lines, representing the motor with the nonlinear rate feedback, can be replaced by a describing function $G(j\omega, \hat{\theta}_m)$ given by

$$G(j\omega,\hat{\theta}_m)=\frac{K}{j\omega(j\omega T+1)}\,\frac{1}{1+\frac{kKN(\hat{\theta}_m)}{j\omega T+1}}$$

$$=\frac{K}{j\omega[j\omega T+1+KkN(\hat{\theta}_m)]} \qquad (1)$$

where K is the loop-gain, k is the feedback loop-gain, T is the motor time constant, and $N(\hat{\theta}_m)$ is the describing function of the block N and is given by



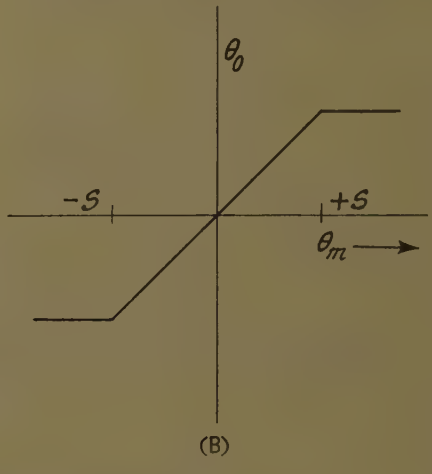


Fig. 5. A—Schematic of a nonlinear rate-feedback system. B—Characteristics of block N

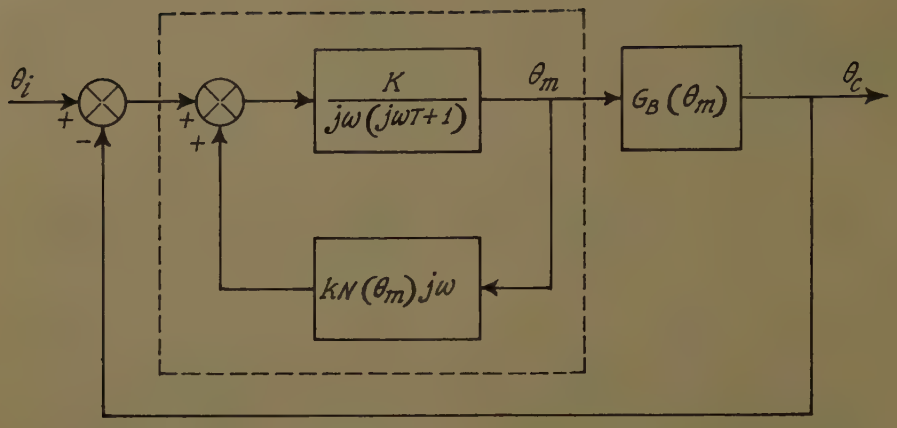


Fig. 6. Block diagram of system in Fig. 5 (A), using frequency-response functions

$$N(\hat{\theta}_m)=\frac{2}{\pi}\left[\sin^{-1}\frac{s}{\hat{\theta}_m}+\frac{s}{\hat{\theta}_m}\sqrt{1-\frac{s^2}{\hat{\theta}_m{}^2}}\right] \qquad (2)$$

The simplified equivalent system can then be represented as in Fig. 7, and the condition for sustained oscillations is

$$G(j\omega,\hat{\theta}_m)=-[G_B(\hat{\theta}_m)]^{-1} \qquad (3)$$

$G_B(\hat{\theta}_m)$ being the describing function of the backlash characteristics.

Validity of this condition in any system is tested most conveniently in a graphical manner, which also indicates suitable means for avoiding oscillations. The two functions in equation 3 are superposed, either in the complex (Nyquist or inverse Nyquist) plane or in the gain-phase shift plane. In Fig. 8, for example, are plots of these functions for assumed values of $K=10$, $k=0.1$, and $T=1$; n and m are normalized amplitude parameters, being respectively equal to $\delta/\hat{\theta}_m$ and $S/\hat{\theta}_m$.

While the plot of $-G_B(\hat{\theta}_m)^{-1}$ is an amplitude-dependent locus, those of $G(j\omega, \hat{\theta}_m)$ are a family of frequency-dependent loci, each locus being drawn for a specific signal amplitude. Thus, validity of relation 3 requires that there be an interesection between the $-G_B(\hat{\theta}_m)^{-1}$ locus and the particular $G(j\omega, \hat{\theta}_m)$ locus, which bears the same amplitude mark as the $-G_B(\hat{\theta}_m)^{-1}$ locus at the intersection point.

The forms of the $G(j\omega, \hat{\theta}_m)$ loci in Fig. 8 reveal that, as a result of nonlinear rate feedback, the frequency-response transfer function of the motor unit has comparatively less gain and less lagging phase shifts at small-signal levels. This indicates a small-signal stabilizing effect of the nonlinear-rate feedback scheme of Fig. 5(A), and points out the possibility of avoiding sustained oscillations in the system by proper choice of the parameters of block N.

Compensation for Hysteresis Effects

Since the describing functions $G(j\omega, \hat{\theta}_m)$ and $G_B(\hat{\theta}_m)$ have been derived for $\hat{\theta}_m(t)$, which are assumed sinusoidal, the input signal $\theta_i(t)$ is assumed to be a periodic function, making $\hat{\theta}_m(t)$ vary sinusoidally. The relations then are as follows:

$$\theta_m/\theta_e=G(j,\hat{\theta}_m)\omega \qquad (4$$

$$\theta_c/\theta_m=G_B(\hat{\theta}_m) \qquad (5$$

The total forward-loop transference i now

$$\theta_c/\theta_e=G(j\omega,\hat{\theta}_m).G_B(\hat{\theta}_m) \qquad (6$$

and the closed-loop transference is

$$\theta_c/\theta_i=G_B(\hat{\theta}_m)/[G(j\omega,\hat{\theta}_m)^{-1}+G_B(\hat{\theta}_m)] \qquad (7$$

The frequency-response curves repre sented by equation 7 can be computed wit the signal amplitude as a parameter b following a graphical procedure.[10]

Proper compensation for hysteresis effects requires that $N(\hat{\theta}_m)$ be chosen i such manner that the amplitude de pendence of these curves will be reduce to a minimum. This can be done by tria and error. But to design a compensate system by this procedure proves tediou and impractical. Adjustments are bes made on the basis of computer runs an tests of a bread-boarded system.

Nonlinear Phase-Lead Compensation

General Facts

The linear phase-lead network show in Fig. 9(A) has the gain and phase fun tions shown in Fig. 9(B). Parameters the network generally are so adjusted tha the maximum phase lead occurs near th crossover frequency, resulting in improve phase margin with a consequent increas of relative stability. If such a netwo

Fig. 7 (below). Equivalent representation of system in Fig. 6 for stability analysis

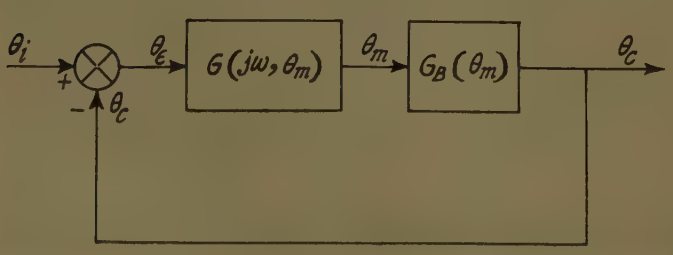
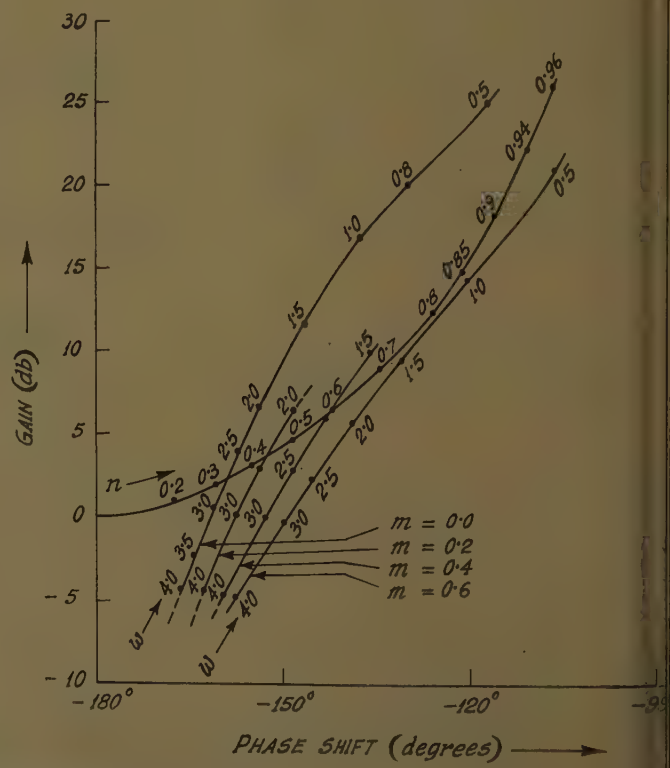
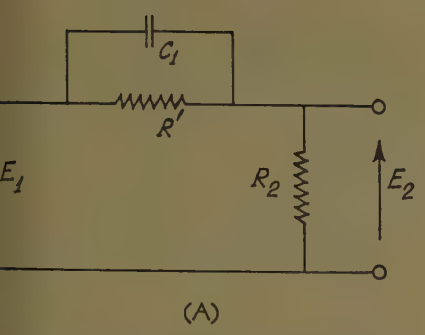


Fig. 8 (right). Nichols plots for system in Fig. 7 for K=10 and T=1





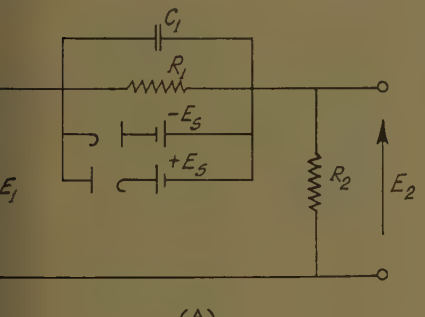
(A)

g. 9. A—RC phase-lead network. B—ots of transfer gain and phase of the network

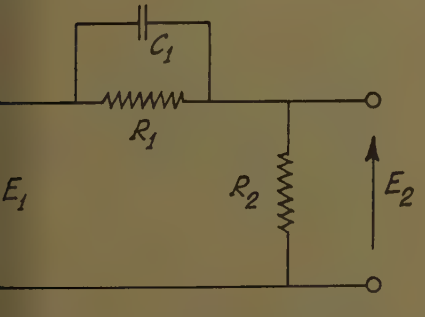
used to compensate for the phase-lag fects of hysteresis, the resultant system ith greater phase margins at comparavely large signals implies that the system sponse should be somewhat sluggish for rge-signal amplitudes. This apparent sadvantage is easily remedied by nonnearizing the network depicted in Fig. (A).

Nonlinear Phase-Lead Network

The network shown in Fig. 10(A) ffers from the one in Fig. 9(A) in that vo biased diodes have been added. The orking of this circuit can be explained ith the help of the equivalent circuit own in Fig. 10(B) where R' is a nonlinear sistor, equivalent to the parallel comnation of R_2 and the diode conduction sistance. At small-signal levels, where $E_s<E_1-E_2<E_s$, both the diodes will be onconducting and $R'=R_2$. The network en behaves in the usual manner. If, owever, the signal amplitude is large ough to reverse the inequality, one

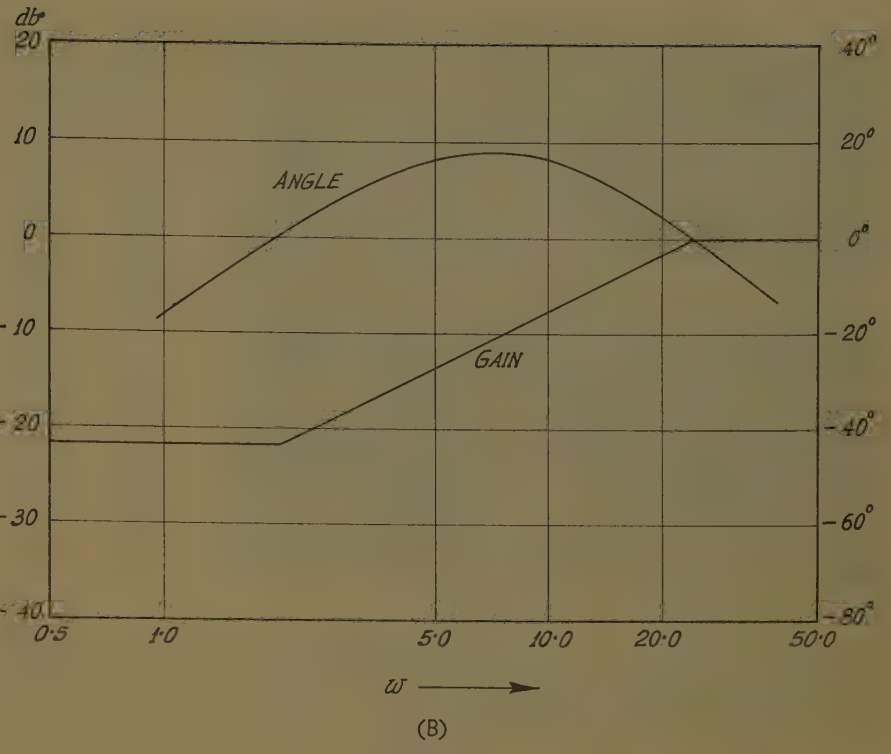


(A)



(B)

g. 10. A—Proposed nonlinear-lead network. B—Equivalent of the network



(B)

diode conducts, depending on signal polarity. Since diode conduction resistance is extremely small compared with R_2, R' is also quite small and the circuit effectively behaves as a unity-transference network. The simple network of Fig. 10(A) thus is seen to introduce amplitude-dependent phase leads, which predominate at small-signal amplitudes only.

Effects of Nonlinear Lead Network

The effects of incorporating the network of Fig. 10(A) in a nonlinear feedback system with hysteresis can be discussed on the basis of network frequency response. In Fig. 11, the hysteresis derives from backlash effects. The lead network is placed in tandem in the forward-loop after the error-sensing unit. This unit has a quasilinear frequency-response function $G_C(j\omega, \hat{\theta}_\epsilon)$, as indicated, which is a function of both frequency ω and signal amplitude $\hat{\theta}_\epsilon$.

A derivation of this transfer function is difficult because conduction of the diodes depends on the signal difference E_1-E_2 $(\theta_1-\theta_0)$ rather than on E_1 $(\hat{\theta}_\epsilon)$ alone. (An easily workable, though approximate, describing function for the network is derived in the Appendix.)

$$G_c(j\omega,\hat{\theta}_\epsilon)=[A_1^2+B_1^2]^{1/2}\ \underline{/\tan^{-1}(B_1/A_1)} \quad (8)$$

$$A_1=\left[q/p+\frac{2}{\pi}(1-q/p)\cos^{-1}\frac{E_s}{\hat{E}_1}-\frac{2}{\pi}\frac{E_s}{\hat{E}_1}(1+q/p)\sqrt{1-\frac{E_1^2}{\hat{E}_1^2}}\right] \quad (9)$$

$$B_1=\left[\frac{\omega}{\pi p}\left(\pi-2\cos^{-1}\frac{E_s}{\hat{E}_1}+2\frac{E_s}{\hat{E}_1}\times\sqrt{1-\frac{E_s^2}{E_1^2}}\right)-\frac{4q}{\pi p}\frac{E_s^2}{\hat{E}_1^2}\right] \quad (10)$$

The best procedure under the circumstances is an experimental determination of transfer gain and phase as a function of signal frequency and amplitude.

Table I gives values of the instant network's transfer gain and phase for three different values of signal ratio $E_1/\hat{E}_1$, obtained on the basis of equation 8 and by direct measurement, the latter being within about 2% of absolute accuracy.[11] The order of approximation involved in equation 8 can now be estimated.

Function $G_C(i\omega, \hat{\theta}_\epsilon)$ cannot be used as was $G(j\omega, \hat{\theta}_m)$ for determining the limit cycle parameters, since the former is a function of $\hat{\theta}_\epsilon$ while $G_B(\hat{\theta}_m)$ is a function of $\hat{\theta}_m$. A trial-and-error procedure is permissible; for, under conditions of sus-

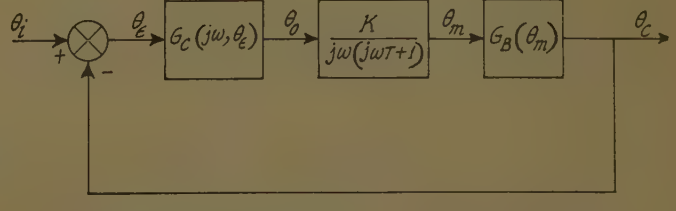


Fig. 11. Block diagram of the nonlinear lead-network compensated system

Table I. Transfer Gain and Angle of Nonlinear Lead Network

Three Regions In Network	Frequency, Radians Per Second	Gain		Angle, Degrees	
		Calculated	Measured	Calculated	Measured
I. $E_s/\hat{E}_1=1.0$	1.86	0.114	0.115	38.5	38.5
	2.06	0.120	0.119	40.9	40.95
	2.62	0.136	0.139	46.3	46.4
	3.02	0.150	0.142	49.2	49.5
	3.59	0.169	0.173	52.4	52.4
	4.40	0.198	0.203	55.2	54.3
	5.70	0.245	0.250	57.4	56.4
	7.85	0.321	0.329	57.6	56.7
II. $E_s/\hat{E}_1=0.8$	2.06	0.128	0.094	5.4	18.2
	2.62	0.131	0.096	13.9	22.3
	3.02	0.137	0.100	21.1	25.4
	3.59	0.145	0.104	28.3	28.8
	4.40	0.162	0.112	37.7	33.4
	5.70	0.194	0.125	49.1	39.4
	7.85	0.264	0.149	61.0	48.0
III. $E_s/\hat{E}_1=0.4$	2.06	0.508	0.3856	3.0	3.05
	2.62	0.509	0.3870	4.2	3.8
	3.02	0.510	0.3872	5.3	4.5
	3.60	0.511	0.3875	6.5	5.3
	4.4	0.513	0.388	8.4	6.3
	5.7	0.518	0.390	11.3	8.4

Table II. Amplitude and Frequency of Oscillation in Nonlinear Rate Feedback System

K	Limiting Amplitude Compared with Backlash (ϵ/δ)	Values of $n=\delta/\hat{\theta}_m$		Values of Frequency, Radians Per Second	
		Simulator	Describing Function	Simulator	Describing Function
10	0.0	0.316	0.315	2.73	2.80
	0.5	0.423	0.440	2.50	2.30
	1.0	0.638	No oscillations	1.96	No oscillations
4	0.0	0.601	0.495	1.3	1.20
	0.5	0.732	No oscillations	1.02	No oscillations

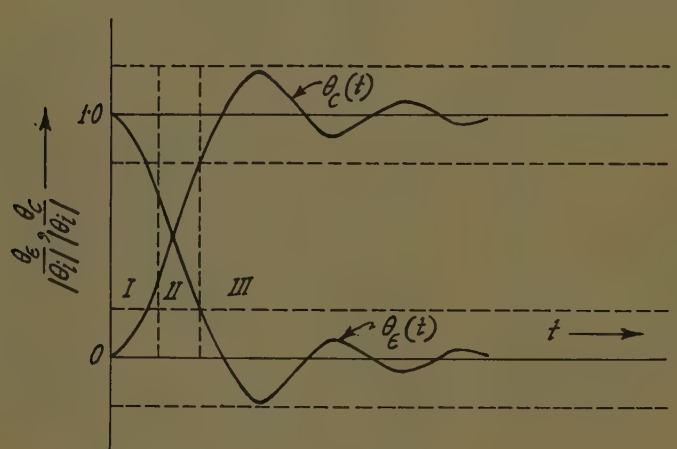
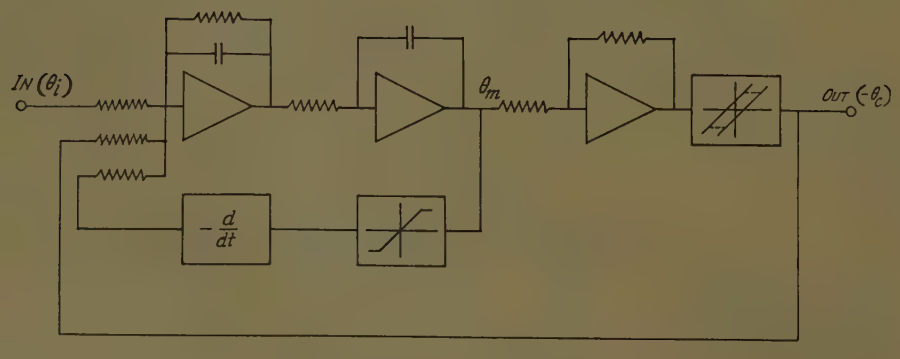


Fig. 12. Typical error and response functions of system in Fig. 11, following step input disturbance



(A)

tained oscillations, the assumption is t
$\hat{\theta}_1(t)=0$. Then $\hat{\theta}_1(t)=-\hat{\theta}_c(t)=-G_B($
$\hat{\theta}_m(t)$, and the condition of amplitude h
ance can be ascertained. However, s
tedium is not necessary to gain und
standing of the nonlinear network's effe

Fig. 12 shows the form of typi
error and response functions of the syst
in Fig. 11 when $\hat{\theta}_i(t)$ is a step displa
ment. The broken lines show the bou
of diode conduction. Network ope
tion can then be explained for three dif
ent regions. Within the broken lin
the network behaves as if diodes w
absent and produces stabilizing pha
lead effects of a conventional linear n
work. This is marked in Fig. 12 as z
III. For slightly larger signals, zone
in Fig. 12, the network introduces ph
leads only slightly less than those in z
III, but its transfer gain is appreciab
reduced (see data for $E_s/E_1=0.8$
Table I). This implies that the stabl
ing action will here be greater for
network than for the conventional lin
network. For even larger signals, z
I in Fig. 12, the network behaves alm
as a unity-transference device and c
not affect system operation.

Conclusions to be drawn from the f
going observations are that, during
initial stages of the system step respo
little difference exists in the two
works when high-frequency transfere
of unity is considered. The nonlin
network should give better response du
later stages and should produce a gre
stabilizing action. These conclusions
checked in the next section against
results of analog computer studies.

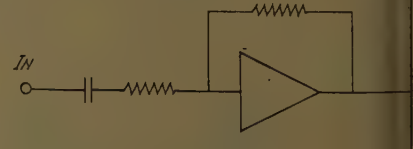
Results of Simulator Studies

Experiments made with the help
analog simulators (real time) verified
preceding results, as described herewit

Nonlinear Rate Feedback System

The simulator is arranged as show
Fig. 13(A). The differentiator unit
in Fig. 13(B) serves reasonably wel
the frequency range of interest. Va
of RC (resistance capacitance) elem
are chosen to yield the transfer func

Fig. 13. A—Simulator arrangement for n linear rate feedback system. (B)—The d rivative unit



(B)

$-10j\omega/(j\omega+100)\cong-0.1j\omega$ for $j\omega<<100$. First, the loop-gain is set high enough o yield sustained oscillations, the rate gnal being set at zero. By fixing the aturation levels at different values, everal simulated rate-feedback systems ere made available for study. Table II ontains amplitudes and frequencies of scillations under different conditions. alues of these parameters, determined y using the describing functions, are also ecorded, and they indicated the approxnate accuracy to be expected of the escribing functions.

The oscillograms in Fig. 14 display step esponse of the system for different setngs of the limiter. The amplitude of ie input step was set at $\theta_i/\delta=10$. The niter greatly increases the speed of sponse, while only slightly increasing ie overshoot. The initial delay in traces Figs. 14(B) and 14(C) is expected, ace full-rate signal is fed back during e initial part of the transient.

onlinear Phase-Lead Systems

The simulator setup is illustrated in g. 15, the compensating network being e one shown in Fig. 10(A). The trans- function of the corresponding linear twork—i.e., Fig. 10(A) with the diodes itted—is $(j\omega+2)/(j\omega+24)$; thus, the aximum phase shift occurs at about rad (radians). Step responses of the stem have been recorded, using two ferent values of the loop-gain: $K=40$ d $K=16$. They give crossover freencies of about 6.3 rad and 4.0 rad, spectively. Figs. 16(A) and (D) show e traces resulting from linear compensan in the two cases. Traces 16(B) and) result from nonlinearization of the mpensation network (with the trace (A) superimposed by way of compari) for E_s/θ_i equal to 0.25 and 0.5, respecely, loop-gain being 40. When the ue of the gain is 16, the traces are as wn in 16(E) and (F).

nclusions

To compensate for their undesirable cts on system stability, the hysterenonlinearities require nonlinear techues. Usual rate-feedback and passive-l networks can be modified easily to p solve hysteresis problems. If a sysi incorporates a tachometer-feedback angement to achieve sufficient stability he presumably linear system, then a linear-gain element (with larger gain maller signals) may be used in the feedk loop to guard against hysteresis cts. Also nonlinearization of the contional RC phase-lead network, by the addition of biased diodes or otherwise, results in better response from the compensated system.

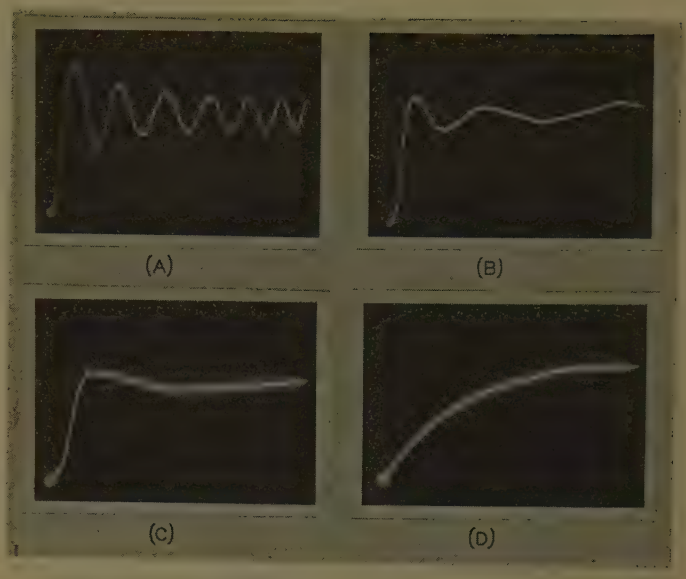


Fig. 14. Traces of step response obtained with setup of Fig. 13(A) where θ_i/δ = 10, K = 10. A—S/δ = 0. B—S/δ = 1. C—S/δ = 2. D—S/δ = ∞

Appendix

Generalized Frequency-Response Function

When diodes are dismounted from the network seen in Fig. 10(A), the transfer function is

$$\frac{E_2}{E_1}(j\omega)=\frac{j\omega+q}{j\omega+p}=\frac{j\omega+2}{j\omega+24} \tag{11}$$

If the diodes are replaced, one of them will conduct for $/E_1-E_2/\ E_s$. If starting with $E_1=E_1\cos\omega t$ so that at $t=0$, the lower diode with $+E_s$ on the cathode will be conducting, and

$$E_2(t)=E_1(t)-E_s \tag{12}$$

the capacitor C_1 being charged to $+E_s$. This continues until $E_2(t)=0$ at $t=t_1$. After this instant, the effective circuit is the one shown in Fig. 17(A) which differs from the conventional linear-lead network by having the condenser charged to $+E_s$. The output voltage, after $t=t_1$ is given by

$$E_2(t)=[E_1(t)-E_s]\left(\frac{s+q}{s+p}\right)\quad\left[s=\frac{\alpha}{dt}\right] \tag{13}$$

This will continue until a time $t=t_2$, perhaps, when $E_1(t)-E_2(t)$ become equal to $-E_s$, and the upper diode starts conduction. At $t=t_2$

$$E_2(t)=E_1(t)+E_s \tag{14}$$

The switching instants t_1 and t_2 are illustrated in Fig. 17(B). The first instant is given by

$$\omega t_1=\cos^{-1}\left(\frac{E_s}{E_1}\right) \tag{15}$$

To find the other instant, some approximations are made. Since $p>>q$, the relationship, taken from equation 13, is

$$E_2(t)\simeq\frac{E_1(t)}{p}(s+q)-\frac{E_sq}{p}$$

or from equation 14

$$E_2(t_2)-E_1(t_1)=\frac{E_1(t_2)}{p}(s+q-p)-\frac{q}{p}E_s=E_s$$

Thus

$$\frac{E_1(t_2)}{p}(s+q-p)=E_s(1+q/p)$$

or $\omega\hat{E}_1\sin\omega t_2-(p-q)\cos\omega t_2=E_s(p+q)$

This gives for t_2

$$\omega t_2=\cos^{-1}\left[-\frac{E_s}{\hat{E}_1}\frac{p+q}{[(p-q)^2+\omega^2]^{1/2}}\right]-\tan^{-1}\frac{\omega}{p-q}$$

$$\simeq\pi-\cos^{-1}\frac{E_s}{E_1} \tag{16}$$

at low frequencies.

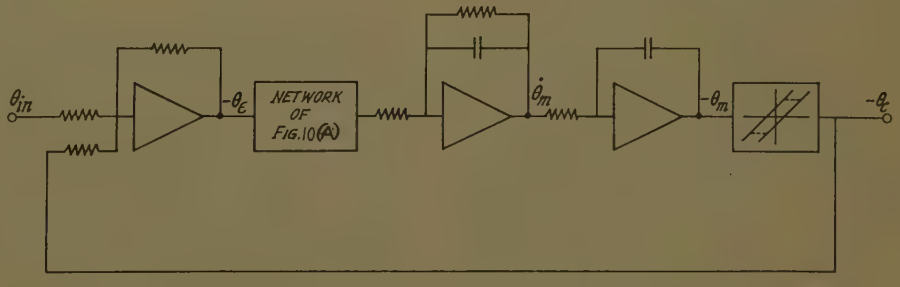


Fig. 15. Simulator arrangement for the nonlinear lead-network compensated system

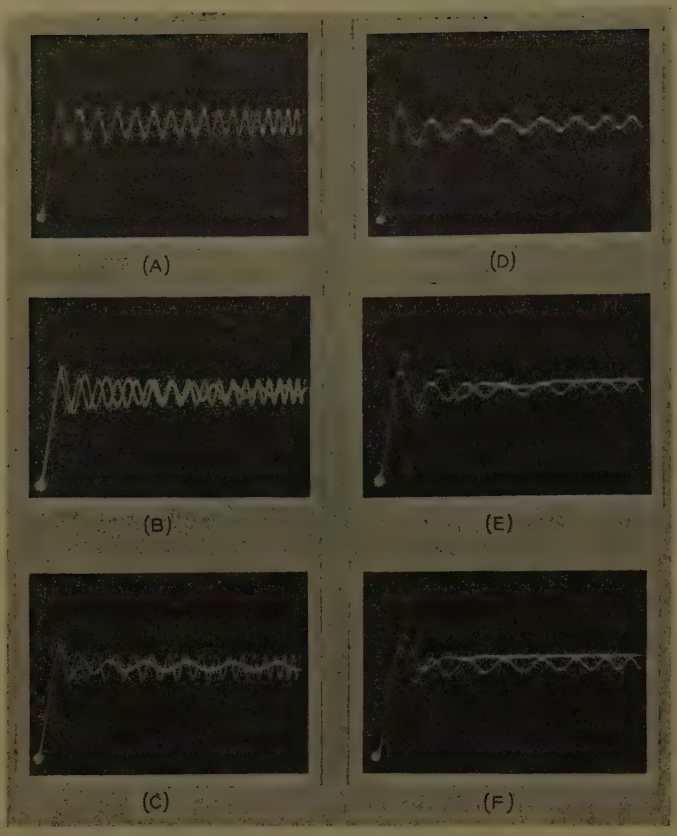


Fig. 16. Traces of step responses obtained with the setup of Fig. 15, where $\theta_i/ = \delta$ 10. A—K = 40, $E_s/\theta_i > 1$. B—K = 40, $E_s/\theta_i = 0.25$. C—K = 40, $E/_s\theta_i =$ 0.5. D—K = 16, $E_s/\theta_i > 1$. E—K = 16, $E_s/\theta_i = 0.25$. F—K = 16, $E_s/\theta_i =$ 0.5

The network's generalized frequency-response function is now computed with the help of relations expressed in equations 12 through 16, giving

$$E_1 = \hat{E}_1 \cos \omega t \tag{17}$$

$$E_2 = \hat{E}_1 \cos \omega t - E_s, \quad 0 \leqslant \omega t < \omega t_1 \tag{18A}$$

$$= \frac{\omega \hat{E}_1 \sin \omega t}{p} + \frac{q}{p}(\hat{E}_1 \cos \omega t - E_s), \quad \omega t_1 \leqslant \omega t \leqslant \omega t_2 \tag{18B}$$

$$= \hat{E}_1 \cos \omega t + E_s \quad \omega t_2 \leqslant \omega t \leqslant \pi \tag{18C}$$

Assuming

$$E_2/\hat{E}_1 = A_1 \cos \omega t_1 + B_1 \sin \omega t + A_3 \cos 3\omega t + B_3 \sin 3\omega t + \ldots \tag{19}$$

Then

$$A_1 = \frac{2}{\pi \hat{E}_1}\left[\int_0^{\omega t_1} (\hat{E}_1 \cos \omega t - E_s) \cos \omega t \times d(\omega t) + \int_{\omega t_1}^{\omega t_2} \left\{(\omega/p)\hat{E}_1 \sin \omega t + \frac{q}{p}(\hat{E}_1 \cos \omega t - E_s)\right\} \cos \omega t \alpha(\omega t) + \int_{\omega t_2}^{\pi} (\hat{E}_1 \cos \omega t + E_s) \cos \omega t d(\omega t)\right] \tag{20}$$

$$B_1 = \frac{2}{\pi \hat{E}_1}\left[\int_0^{\omega t_1} (\hat{E}_1 \cos \omega t - E_s) \sin \omega t \alpha(\omega t) + \int_{\omega t_1}^{\omega t_2} \left\{(\omega/p)\hat{E}_1 \sin \omega t + \frac{q}{p} \times (\hat{E}_1 \cos \omega t - E_s)\right\} \sin \omega t \alpha(\omega t) + \int_{\omega t_2}^{\pi} (\hat{E}_1 \cos \omega t + E_s) \sin \omega t \alpha(\omega t)\right] \tag{21}$$

On carrying out the integrations

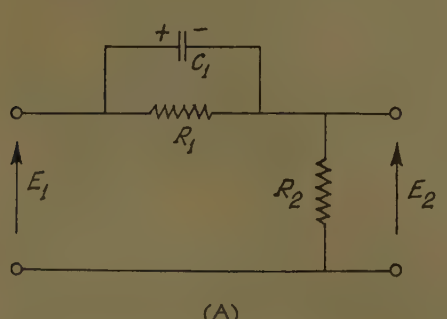
$$A_1 = \left[q/p + \frac{2}{\pi}(1 - q/p) \cos^{-1} \frac{E_s}{\hat{E}_1} - \frac{2}{\pi}(1 + q/p) \frac{E_s}{\hat{E}_1} \sqrt{1 - \frac{E_s^2}{\hat{E}_1^2}}\right] \tag{22}$$

$$B_1 = \left[\frac{\omega}{\pi p}\left(\pi - 2 \cos^{-1} \frac{E_s}{\hat{E}_1} + 2 \frac{E_s}{\hat{E}_1} \sqrt{1 - \frac{E_1^2}{\hat{E}_s^2}}\right) - \frac{4q}{\pi p} \frac{E_s^2}{\hat{E}_1^2}\right] \tag{23}$$

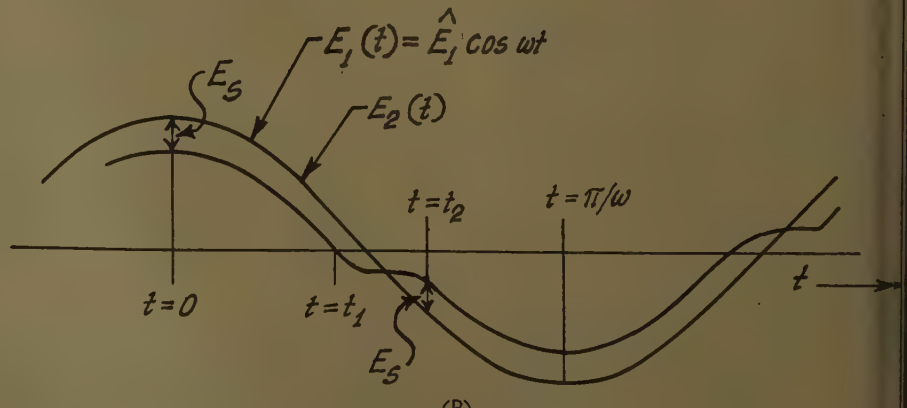
Nomenclature

A_1 = real part of $G_C(j\omega, \hat{\theta}_\epsilon)$
B_1 = imaginary part of $G_C(j\omega, \hat{\theta}_\epsilon)$
C_1 = capacitor
D = width of magnetic hysteresis
E_1 = lead-network input signal
E_2 = lead-network output signal
E_S = diode bias
$G(j\omega, \hat{\theta}_m)$ = describing function of motor wit nonlinear rate feedback
$G_C(j\omega, \hat{\theta}_\epsilon)$ = describing function of nonlinea lead network
$G_B(\hat{\theta}_m)$ = describing function of backlas hysteresis
$G_D(\hat{\theta}_\epsilon)$ = describing function of contactor
$G_H(\hat{\theta}_\epsilon)$ = describing functions of magneti hysteresis
H = nonlinear rate signal
h = relay hysteresis
K = forward loop gain
k = feedback loop gain
$m = S/\hat{\theta}_m$
N = nonlinear block
$N(\hat{\theta}_m)$ = describing function of block N
$n = \delta/\hat{\theta}_m$
p = pole of RC network
q = zero of RC network
R_1, R_2, R' = resistances
$s = d/dt$
S = limiting signal level
T = motor time-constant
δ = backlash width
Δ = dead zone
θ_c = output signal
θ_i = input signal
θ_ϵ = error signal
θ_m = motor output signal
θ_0 = controller output signal
$\hat{\theta}_m$ = amplitude of $\theta_m(t)$ assumed sinusoidal
$\hat{\theta}_\epsilon$ = amplitude of $\theta_\epsilon(t)$ assumed sinusoidal
θ_{es} = saturation level of magnetic hysteres
$\rho = D/\hat{\theta}_\epsilon$
$\lambda = \theta_{es}/\hat{\theta}_\epsilon$
ω = angular frequency

Fig. 17. A—Equivalent of Fig. 10(A) for $t_1 < t < t_2$. B—Illustrative waveform of the nonlinear lead-network output defining the switching instants t_1 and t_2



(A)



(B)

References

FREQUENCY RESPONSE METHOD FOR ANALYZ-ND SYNTHESIZING CONTACTOR SERVOMECHA-R. J. Kochenburger. *AIEE Transactions*, , pt. I, 1950, pp. 270–84.

NUSOIDAL ANALYSIS OF FEEDBACK-CONTROL MS CONTAINING NONLINEAR ELEMENTS, Johnson. *Ibid.*, pt. II (*Applications and ry*), vol. 71, July 1952, pp. 169–81.

EDBACK IN CONTOURING CONTROL SYSTEMS, Ellert. *Ibid.*, vol. 74, 1955 (Jan. 1956 section, 5–54.

ONLINEAR INTEGRAL COMPENSATION OF A CITY-LAG SERVOMECHANISM WITH BACKLASH, Shen, H. A. Miller, N. B. Nichols. *Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 79, 1957, pp. 585–92.

5. BACKLASH COMPENSATION IMPROVES SERVO PERFORMANCE, D. C. McDonald. *Instruments and Automation*, Pontiac, Ill., vol. 28, Nov. 1959, pp. 1728–31.

6. EFFECTS OF COULOMB FRICTION ON THE PERFORMANCE OF A SERVOMECHANISM HAVING BACKLASH, PART II, TRANSIENT RESPONSE CONSIDERATIONS, A. K. Mahalanabis. *Indian Journal of Physics*, Calcutta, India, vol. 35, no. 1, 1961, pp. 1–15.

7. STABILISATION OF CONTACTOR SERVOS BY USING COULOMB FRICTION, A. K. Mahalanabis. *Journal of Electronics and Control*, London, England, vol. 8, no. 4, 1960, p. 307.

8. NONLINEAR COMPENSATION NETWORKS FOR FEEDBACK SYSTEMS, E. Mishkin, J. G. Truxal. *National Convention Record*, Institute of Radio Engineers, New York, N. Y., pt. 4, 1957, pp. 3–7.

9. A DESCRIBING FUNCTION STUDY OF HYSTERESIS EFFECTS IN A VELOCITY-LAG SERVOMECHANISM, A. K. Mahalanabis. *Proceedings*, First Defence Electronics Convention, Bangalore, India, Sept. 28–Oct. 1, 1959, pp. 99–103.

10. FREQUENCY RESPONSE OF NONLINEAR CLOSED-LOOP FEEDBACK CONTROL SYSTEMS, G. H. Fett, S. L. Mikhail. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 77, Nov. 1958, pp. 436–38.

11. A TRANSFER FUNCTION ANALYSER FOR LINEAR AND NONLINEAR COMPONENTS, A. K. Choudhury, M. S. Basu, A. K. Mahalanabis. *Electronic Engineering*, London, England, vol. 33, June 1961, pp. 382–85.

Thermoelectricity Application Considerations

A. A. SORENSEN
MEMBER AIEE

OWADAYS, A DESIGNER is faced with myriad electric power requirets—from milliwatts to megawatts—ed by new weapon and space systems epts. At his disposal, fortunately, umerous and varied potential sources ectricity. There is renewed interest activity in fuel cells and thermionic thermoelectric converters, for exam- These, together with improved bats and rotating generators driven by rted prime movers, present many posities that want consideration for a n system. The particular selection, ourse, is governed by an optimization y, covering required power level, ion time, and desired characteristics. he thermopile is a contender for many ications, now that solid-state physics ies have brought about more efficient uniform thermoelectric materials. It both advantages and disadvantages. ompared with other energy-converdevices, it is simple in construction, e being no electrical fluids or exely close spacings to maintain. g static, the thermopile has the possiy of long life. It operates from chemsolar, or nuclear heat sources, is able se waste heat, and adapts to various ient conditions. On the other hand, pecific weight is high, voltage is low, time constants are long.

ven so, electromagnetic machines rate low voltage from the standpoint individual turns. Similarly, in a nopile series, connections are necesto obtain required voltages. Power sistors and other solid-state devices efficiently convert direct current to desired utilization levels and frequencies. Employing a storage device, long time constants can be handled more easily with d-c than with a-c sources.

Engineering Design

Use of stable metals or alloys for thermocouples in measurement and control applications has grown into an important technology. Semiconductors also have been long recognized as powerful thermoelectric materials. Although only a few semiconductors are elemental in composition, there are numerous semiconducting compounds, many of which will give an electrical output when subjected to a temperature gradient. Some of these are particularly interesting because, having low thermal conductivity, they are economical in their use of heat. General requirements for practical use of these materials in generators are: high electrical output, low thermal conductivity, stability, and adaptability to production processes.

An important phase in any semiconductor materials program is control of the electron or hole concentration in the crystal. In elementary types, such as are found in diodes and transistors, this control is accomplished by adding the desired impurity atoms, which introduce donor or acceptor levels in the semiconductor's band structure. The situation is more complicated in binary compound semiconductors because the dominating carrier type is influenced by the deviation of the compound's composition from stoichiometry. Thus, the material's properties are influenced by both composition and doping agent as well as the method of preparation.

One such thermoelectric material is lead telluride. Couples of the highest efficiency are obtained if *n*-type material is prepared with excess lead over the stoichiometric ratio and *p*-type material with excess tellurium. However, since material with excess lead is mechanically superior to excess tellurium material, it lends itself better to mechanical processes necessary in the construction of a generator.

The preparation method also affects mechanical characteristics. For example, pressed and sintered lead telluride is machined more readily than cast material. Pressing and sintering also contribute to preparing graded thermoelectric arms; that is, to greater doping of the hot end of the arm than the cold end. This is used because the temperature depends on the material's properties and results in some increase in efficiency. Characteristics of graded *n*- and *p*-type arms, produced by pressing and sintering, are shown in Figs. 1 and 2.

More than optimum *n*-type and *p*-type thermoelectric arms are required in engineering applications, however. Also required are a heat source, electrical insulation for hot and cold sides of the couples, thermal insulation, electrical connectors for hot and cold sides of the arms, and a heat sink. As far as the heat source and sink are concerned, the problem in engineering design is to match available heat fluxes with those required for thermoelectric elements. Probably the most troublesome item is joining electrical connections to the thermoelectric arms. Soldering is generally adequate for

Paper 60-1066, recommended by the AIEE Aero-Space Transportation Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Pacific General Meeting, San Diego, Calif, August 8–12, 1960. Manuscript submitted March 23, 1959; made available for printing May 16, 1961.

A. A. SORENSEN is with The Martin Company, Baltimore, Md.

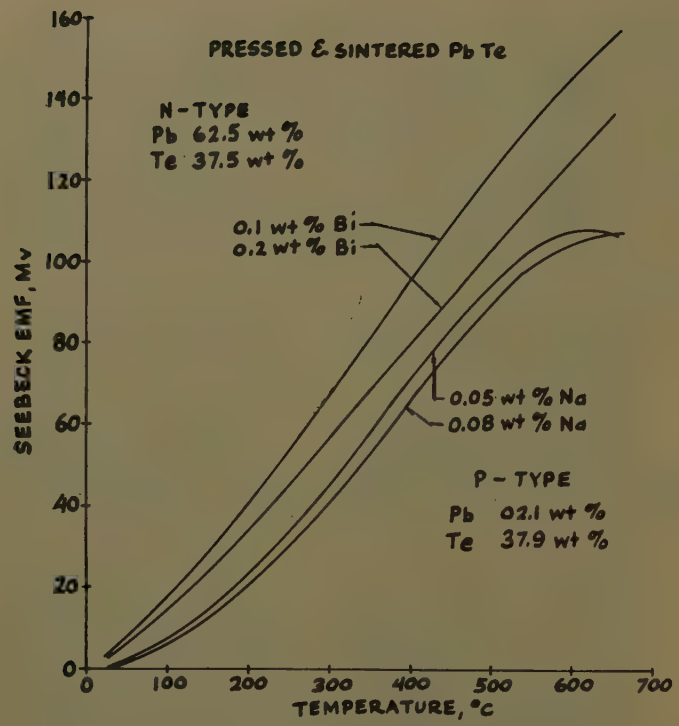


Fig. 1. Seebeck electromotive force versus temperature

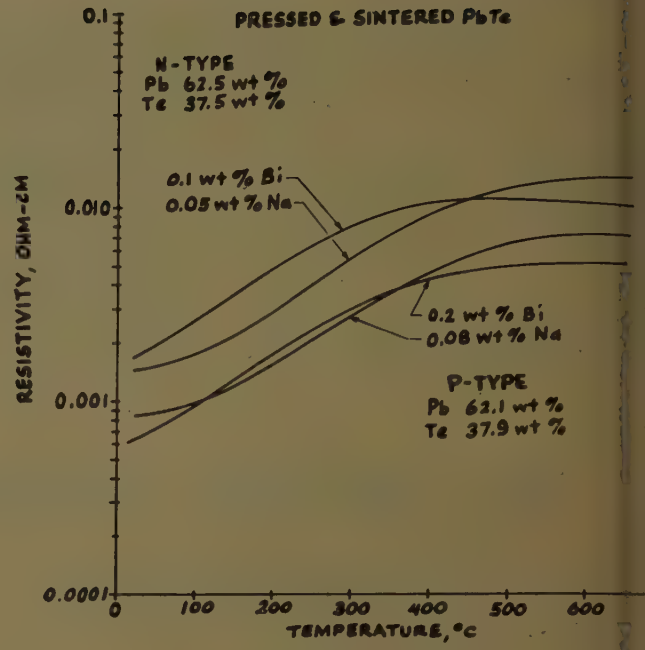


Fig. 2. Resistivity versus temperature

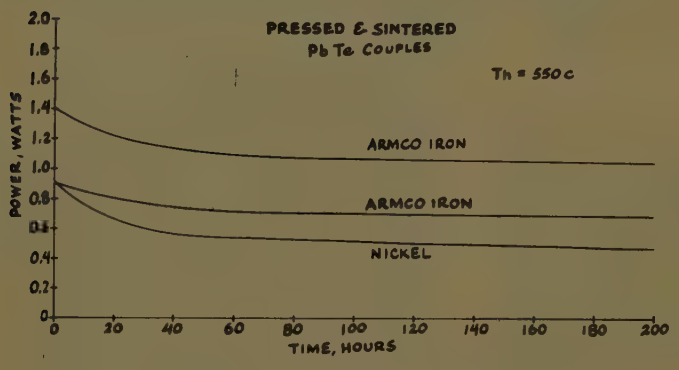


Fig. 3. Power output versus time

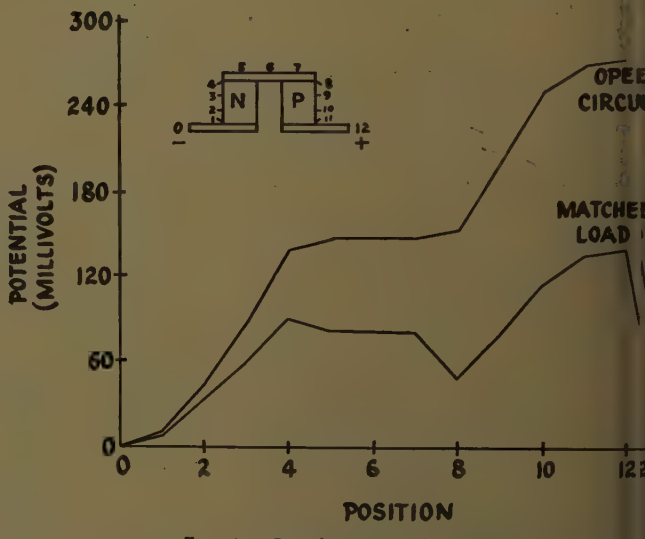


Fig. 4. Couple potential profile

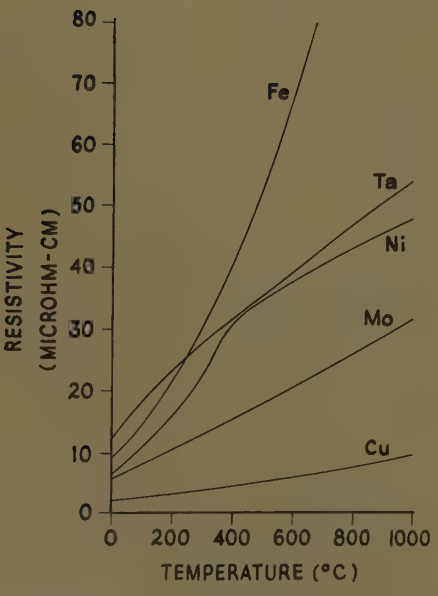


Fig. 5. Resistivity of some metals

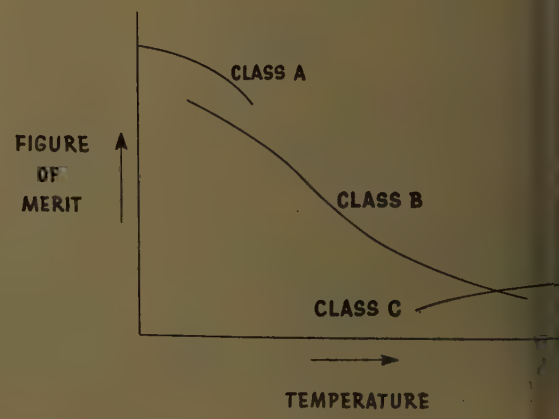


Fig. 6 (right). Figure of merit versus temperature for typical thermoelectric materials

Table I. Space Applications

Applications	Ratings	Limitations
ric power source for rocket stages, ng cryogenic cooling and rocket at	1 to 20 kw (kilowatts) (15 to 25% efficiency)	Weight, thermal log
ric power source for satellites, ng nuclear or solar heat sources	1 watt to 5 kw (5 to 8% efficiency)	Pumping losses limit efficiency intermediate ratings. Heat storage needed for solar source
rol power devices	Milliwatts to watts	

Table II. Terrestrial Applications

Applications	Ratings	Limitations
ulsion for automotive equipment	5 to 10 kw (10% efficiency)	Best suited for low speed, intermittent use. Battery needed for peak loads
bustion-powered, portable electric-wer source for quiet operation	100 watts to 5 kw (5 to 10% efficiency)	Pumping losses an appreciable part of output
r-powered battery chargers for use remote areas	25 watts to 1 kw (2 to 7% efficiency)	Orientation and concentration required for higher efficiencies

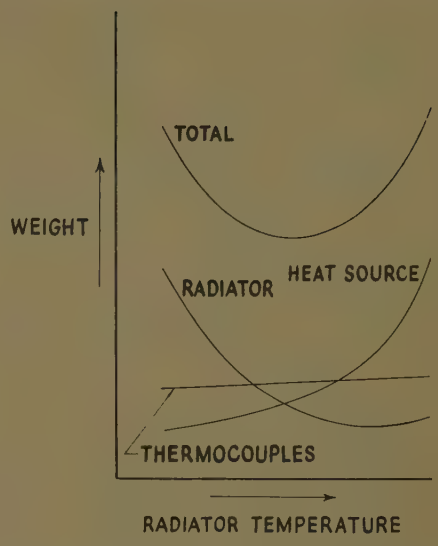


Fig. 7. Graphical solution for minimum weight for heat source, thermocouples, and radiator

cold sides. For the hot sides, how-r, materials must be used which do not ct the performance of either *n*- or *p*-e arms. Fig. 3 shows the output of ee couples made with lead telluride, with iron electrodes, and one with kel. The drop-off in output with time nore rapid for the nickel electrode because the hot end of the *p*-type material changes to *n*-type.

A helpful tool for gauging over-all performance of a thermocouple is the couple-potential profile shown in Fig. 4. Fine wire voltage pickups are attached at points on the thermocouple as indicated by the numbers shown. Voltages across cold-end connections are indicated by 0 to 1 and 11 to 12; across the hot-end connections by 4 to 5 and 7 to 8. The open-circuit profile of the couple shows voltage distribution between *n*- and *p*-type arms. This also reveals deterioration in performance during extended operation and points out the cause of the difficulty, such as reverse doping of an arm by the electrode. In this case, a reverse voltage is seen at the hot end of the arm involved. The matched-load profile permits determination of contact resistances under operating conditions and shows output distribution between the two arms. The profile also allows voltage drop in the body of the electrode connecting the two arms to be checked.

The hot connecting electrode can be-

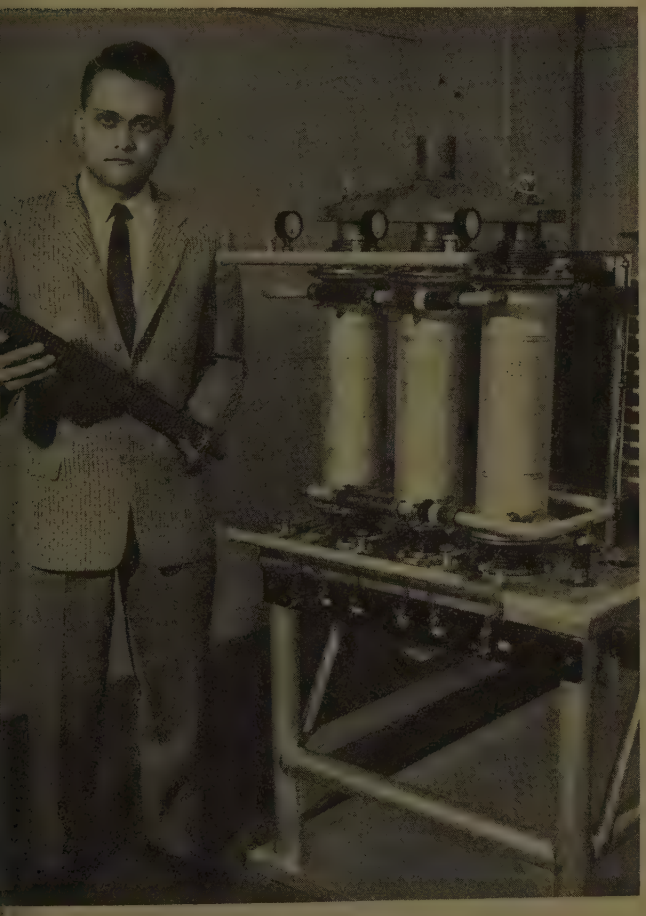


Fig. 8. Thermoelectric generator

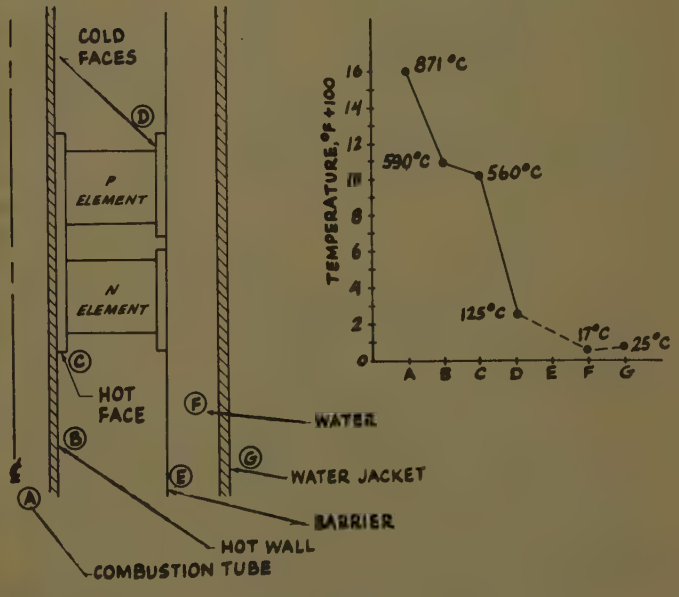


Fig. 9. Temperature profile

Table III. Marine Applications

Applications	Ratings	Limitations
Main propulsion	500 kw and up (15 to 20% efficiency)	Best used with battery for system voltage control
Auxiliary power	250 watts to 100 kw (5 to 20% efficiency)	
Signal devices	5 to 50 watts (10 to 15% efficiency)	

Table IV. Net Thermocouple Efficiency

Applications	Temperature		Efficiency		
	T_h	T_c	Eff_c	Eff_{tc}	Eff_N
Space	2,000 degrees Fahrenheit (1,093 degrees centigrade)	900 degrees Fahrenheit (482 degrees centigrade)	45%	17%	7.7%
Terrestrial	1,600 degrees Fahrenheit (871 degrees centigrade)	150 degrees Fahrenheit (66 degrees centigrade)	70%	26%	18%
Marine	1,600 degrees Fahrenheit (871 degrees centigrade)	45 degrees Fahrenheit (7 degrees centigrade)	75%	28%	21%

come an appreciable weight item, particularly when contact resistance is lowered and the element lengths decreased. Fig. 5 shows resistance versus temperature for some of the metals of interest for contacts. At higher temperatures, composite connectors are indicated to obtain minimum connector volume.

Efficiency is dependent upon the figure of merit (usually designated by Z)—a composite of the Seebeck coefficient, thermal conductivity, and resistivity. Thus far, the highest figures of merit have been attained by materials suitable for use in room-temperature range. The general trend as temperature increases is shown in Fig. 6. Representative of class A are various alloys of Bi_2Te_3. Class B includes PbTe and InSb, while class C covers oxides and silicides. Conditions under which design-work operation is to take place must be established—the heat sink temperature, range of temperature from the heat source, required life, and characteristics required by the load. From the standpoint of effect upon the thermopile, designs fall into three general categories: space, terrestrial, and marine.

Space

From the designer's viewpoint, space is the worst category because radiation is the only heat-dissipating agent available. For small power applications, conduction may be used to transport heat to the radiator. If the power is appreciable, though, weight becomes excessive for this method, and a circulating fluid system must then be used. When a design is optimized for weight by compromising between the radiator's weight and that of the thermopile and heat source, then the radiator operates at a comparatively high temperature. Optimization may be carried out by analytical solution of equations, by a computer, or by graphical solution based on calculations of several cases, as shown in Fig. 7.

Since the temperature gradient in the elements is assumed to be fixed, the weight of the thermoelectric elements will not change greatly with the radiator temperature. Thus, for a fixed maximum, the greater number of elements required at a higher radiator temperature is compensated by their shorter length. This assumption requires that contact resistance must not be the governing factor in element length. Table I shows some possible space applications.

Terrestrial

Although nuclear, solar, and combustion heat sources are available for terrestrial use, the latter likely will provide for the majority of cases. Catalytic combustion processes are especially favorable for thermoelectric generation, since an extended high-temperature area can be provided and heat transfer is largely by radiation, thus reducing stack losses. Fig. 8 shows a generator employing this combustion principle; the catalytic combustion tube held by the young man is

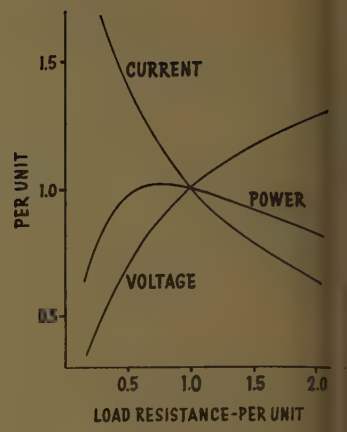


Fig. 10. Output parameters versus loa

similar to ones inside each of the t conversion tubes. At the time the ph graph was taken, the output of generator with pressed and sintered telluride couples was 130 watts. Eac the stacks weighed 46¾ pounds. F shows a profile of temperature distribu from heat source to sink.

For terrestrial uses, as contraste space, both radiation and convec effects will dispose of unusable Thus, radiator weight can be redu However, pumping power is neede move air through a radiator, and a an optimization must be made betw over-all efficiency and weight. Mili specifications for ground application the maximum cooling-air temperatur 125 degrees Fahrenheit, and lowe higher, depending upon the use, for borne classifications. Table II possible terrestrial examples.

Marine

Marine applications, some of which outlined in Table III, are the most fa able from a heat-disposal viewpoint cause water is available in quantity temperature no higher than 80 deg Fahrenheit. Thus, the machinery heat disposal is simple and the pump power requirements are a small centage of the total power generated.

The design of a complete power- version system is governed not only

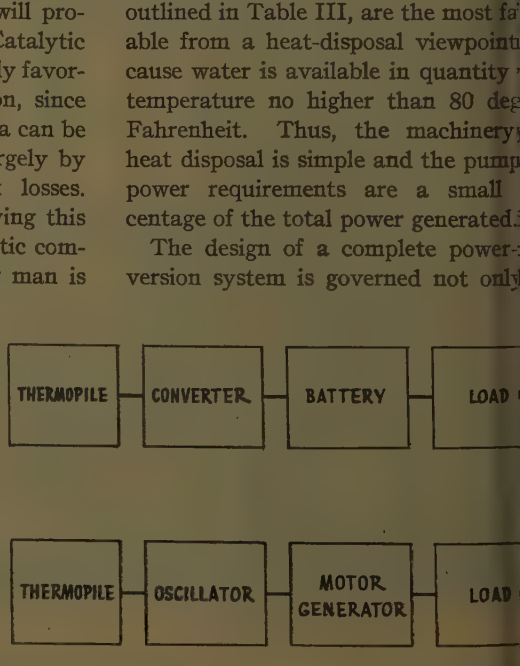


Fig. 11. Arrangements for supplying cyclic loads

apabilities of thermoelectric materi-
but is also strongly influenced by
erature limitations of the heat source,
rodes, seals, support materials, and
rical and thermal insulation. Al-
gh the marine application is the most
rable from the efficiency standpoint,
does not infer that it alone is prac-
In each area, thermoelectric power
ms are subject to the same engineer-
material limitations as competing
ms, except for one advantage: they
not withstand the high stresses
untered with rotating systems. The
tive for operation at higher tempera-
s in space applications will demand
time and effort in materials de-
oment and engineering effort. A pro-
on of near-term thermocouple effi-
ies for the three areas is shown in
e IV. Over-all equipment efficiencies
be lower because of the factors pre-
sly discussed.

Methods of Use

With currently possible efficiencies of about 10%, the value of m (ratio of external to internal resistance) is approximately 1.3. This means that, at the maximum efficiency point, the ratio of open-circuit to closed-circuit voltage with rated load is 1.77, or a regulation of 43.5%, which is equivalent to that of an unregulated a-c generator. With improved materials, m will be larger and the regulation lower in value. Voltage, current, and power are plotted against per-unit resistive load in Fig. 10. Note that power output is nearly constant for a wide range of load resistance. Applications requiring a constant power input, such as some motor drives, could take advantage of this characteristic. In other applications, needing direct current at various voltages, transistor converters with internal voltage regulators may be used. Similarly, for a-c needs, inverters are chosen.

Thermopiles have the important function of supplying continuous power to a storage device, such as a battery or rotating flywheel, from which energy is withdrawn as needed. In this application, the thermopile can be operated at its most efficient load. Intermediate equipment performs the task of regulating the voltage and converting it to the required form. Two arrangements are shown in Fig. 11.

Each source of electric energy—be it an electromagnetic generator, primary battery, or thermopile—has its distinctive characteristics and limitations, which engineers recognize and take into account by fitting designs to the various idiosyncrasies. With the thermopile becoming generally available, engineers will become adept in designing around its limitations and exploiting its advantages.

---

ome Recent Advances in the Analysis and Synthesis of Nonlinear Systems

ALFRED A. WOLF
MEMBER AIEE

URING the past two decades a great deal of material has appeared in the ature concerning the analysis and hesis of nonlinear systems. In refer- 1, a great many of these techniques reviewed and it is quite evident that e methods are highly specialized aining to particular problems. Some linearization techniques and others approximations, sometimes quite icted. Each has its particular value usefulness. Despite the number of rs, there seems to be no underlying nifying theory connecting them.

ery recent publications appearing in iterature[2-8] have shown the develop- t of a mathematical theory for the ysis of a class of nonlinear systems. nning with the author's dissertation,[2] h gives a good account of this theory, ies of articles have appeared expound- the deterministic process of analysis onlinear systems and extending the ry to systems subjected to stochastic esses with the aid of the Statistical- nsform Theorem.[9,10]

ose,[11] working in collaboration with ner[12,13] and Lee,[14] has developed a inear theory from a point of view quite different from that developed by the author in collaboration with Ku and Dietz. The Wiener-Lee-Bose theory emphasizes a statistical approach to the synthesis of nonlinear systems taking advantage of orthogonal networks such as the Laguerre and Hermite systems.[15]

It is the purpose of this paper to develop new theoretical relations between the statistics of the output of a system, when subjected to a white-noise probe, and the configuration of the topology of the system's actual structure. As a direct implication of this last point, a relation exists between theoretical facets of analysis and the corresponding aspects of synthesis. In addition, as a point of departure, the partition theory will be discussed and extended.[2,3] It will be shown that although the solution of a certain class of nonlinear systems is unique, the form of solution is not unique and is a function of the point of partition. Examples will be given to illustrate this. Finally, consideration will be given to a synthesis procedure for specified waveforms.

Other possible approaches to the solution of nonlinear systems will not be considered since they are adequately treated in reference 1.

Partition Theory Analysis

The systematic study of physical nonlinear systems is based on a study of a certain class of nonlinear differential equations. Consider a typical equation of this class:

$$Z(D)x(t)+F\{x(t),\dot{x}(t),\ldots\}=g(t) \qquad (1)$$

where the linear part $Z(D)x(t)$, the nonlinear part $F\{\ldots\}$, and the driving functions $g(t)$, are suitably restricted by the broad class of conditions given in Appendix I. A result of imposing these conditions is that the solution is unique and analytic. Keeping these two important

---

Paper 61-713, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE-AIChE-ASME-IRE-ISA Joint Automatic Control Conference, Boulder, Colo., June 28–30, 1961. Manuscript submitted October 26, 1961; made available for printing May 8, 1961.

ALFRED A. WOLF is with Emertron, Inc. (a subsidiary of Emerson Radio and Phonograph Corporation), Silver Spring, Md.

The author would like to thank Dr. Y. H. Ku of the University of Pennsylvania whose helpful discussions and comments were much appreciated, Dr. L. F. Kazda of the University of Michigan who suggested and encouraged this study, Dr. T. L. Higgins of the University of Wisconsin for his encouraging remarks on this work, Dr. J. Randolph, Dr. D. Shen, Dr. L. Kanal, and Mr. J. Dietz for many hours of stimulating discussion. The author is also grateful to Mr. K. Lord, Dr. R. Weller, and Mr. D. Keim of the Stromberg-Carlson Company for their encouragement, and finally Mr. Parkhill and Mr. Rapuano for their comments. This paper is a modified version of CP60-110 presented at the AIEE Winter General Meeting, New York, N. Y., February 4, 1960.

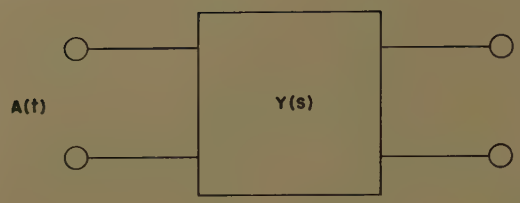


Fig. 1. A linear system with response x(t)

results in mind, consider another problem. Referring to Fig. 1, let $Y(s)$ define a transfer function of a linear system related to the linear differential operator $Z(D)$ according to the relation

$$Y(s)=\frac{1}{Z(s)} \tag{2}$$

This is equivalent to the relation[16]

$$y(t)=\frac{1}{2\pi j}\int_{Br_1}\frac{e^{st}}{Z(s)}\,ds \tag{3}$$

by noting it is the inversion of equation 2 in the complex s-plane along a Bromwich contour (Br_1) enclosing the poles of $Z^{-1}(s)$. An interesting question arises: Does a function $A(t)$ exist such that when it is applied to the input terminals of $Y(s)$, shown in Fig. 1, a response $x(t)$ results identical with the response of a nonlinear system driven by $g(t)$ described in equation 1? The answer, from strictly physical considerations, is evidentally in the affirmative. That this is true is clear at once if the nonlinear system is restricted to physical systems. Then $x(t)$ is physically realizable.[17] Hence, $A(t)$ is physically realizable and therefore it does exist. If function $A(t)$ were known, it would be an easy task to calculate $x(t)$, so that $x(t)$ would be determined by the familiar convolution integral:

$$x(t)=\int_0^t A(\tau)\,y(t-\tau)d\tau \tag{4}$$

Utilizing the conditions of Appendix I, the author showed in references 2 and 3 a rigorous mathematical basis for the existence and uniqueness of $A(t)$, the auxiliary forcing function. Moreover, $A(t)$ is analytic under these conditions.

A hint as to how to calculate $A(t)$ is obtained from physics;[18] consider Fig. 2. This system will be called the canonical system since it can be described by equation 1 which is the general form for a wide variety of physical systems satisfying conditions given in Appendix I. Suppose now that this system is partitioned as shown in Fig. 3 by making cuts at points a and b, shown in Fig. 2. By properly maintaining the system dynamics, the response of the linear part behaves exactly as the response of the interconnected system.

The equations describing the partition in Fig. 3 are

$$Z(D)x(t)=A(t) \tag{5}$$

and

$$A(t)=g(t)-F\{x(t)\} \tag{6}$$

Equation 6 is known as the auxiliary equation[2,3] while equation 5 is called the partition equation. It is evident that these equations could have been derived directly by mathematical manipulation of equation 4 with equation 1. What do these equations mean? Obviously, eliminating $A(t)$ from equations 5 and 6 does not achieve any useful result except to give back equation 1. To use equations 5 and 6 effectively, advantage must be taken of the a priori properties of $A(t)$, which were deduced by means of the conditions required of the linear and nonlinear part of equation 1 and its forcing function. From these conditions, $x(t)$ is deduced to be analytic. Since $g(t)$ is analytic and F is continuous ($A(t)$ is analytic since it is the difference between two analytic functions), the function F is taken to have the form

$$F(x)=\sum_{k=2}^{N} a_k x^k \tag{7}$$

This is not a necessary condition. If F does not have the form of equation 7, it must then be analytic. This condition may, in certain circumstances, be relaxed with due caution.

Thus there is justification for expanding $A(t)$ into a power series. Under these conditions the solution of equation is given by the moment theorem.

Moment Theorem

Given an ordinary nonlinear differen equation (such as equation 1 suita restricted according to the conditi given in Appendix I), the solution $x(t)$ linear combination of all the moments the folded impulse response of the lin part $Z(D)$ where the coefficients are gi as a recurrence equation.[2–5]

Formally, the moment theorem can written as follows. Let $Q_n(t)$ denote nth moment of the folded impulse sponse of the linear part of equation 1 o the half-open interval $(0, t)$, i.e.,

$$Q_n(t)=\int_0^t \tau^n y(t-\tau)d\tau$$

Then the solution of equation 1 is:

$$x(t)=\sum_{n=0}^{\infty} c_n Q_n(t)$$

The coefficients C_n are determined cursively. When $F(x)$ is a polynomial that given by equation 7 and the syst is initially at rest, i.e., all initial conditi are zero, the recurrence equation equation 1 is given by equation 10:

$$C_n=b_n-\sum_{k=2}^{M} C_{n-\alpha}{}^{(k)}a_k;\ \alpha\geq 0 \text{ and}$$

integral valued

where the b_n's are the Taylor coefficie of the expansion of the forcing funct $g(t)$. Equation 10 is obtainable dire by the Taylor-Cauchy transform.[3]

Rules for Partition

1. The dynamical equation represe ing the system is partitioned so that linear terms involving the highest-or

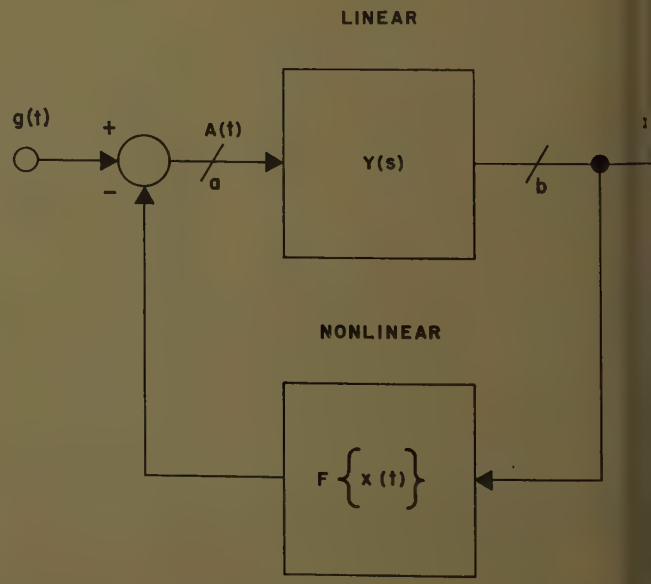


Fig. 2. A canonical nonlinear feedback system

tive are contained in one member
equation.

A modified forcing function $A(t)$ is
btained which consists of the actual
function, nonlinear and possibly
linear terms.

A proper choice is made of the ex-
n of $A(t)$ according to the particular
rties of the linear, nonlinear, and
g functions. An auxiliary equa-
thus formed.

The properly partitioned differen-
uation is then solved.

Certain coefficients which arise
the auxiliary equation are then cal-
d.

These coefficients, which appear in
partitioned equations, are then
ated to obtain the exact solution.

r-Cauchy and Generalized ansforms

en the partition is made such that
ighest-order derivative remains by
in left member, the moment func-
$Q_n(t)$ is a power function. Hence
lution, $x(t)$, is a power series. The
ence relations for this partition are
able by means of the Taylor-
y transform given by equations 11
2:

$$\frac{1}{2\pi j}\int_C \frac{W^{(k)}(\lambda)}{\lambda^{n+1}}\,d\lambda \tag{11}$$

$$\lambda)=\sum_{n=0}^{\infty} w_{n,k}\lambda^n \tag{12}$$

ransform pairs $w_{n,k}$ and $W^{(k)}(\lambda)$ may
lated operationally by the Taylor-
y operator $\mathfrak{I}_c$ according to the
ons

$$\mathfrak{I}_c\{W^{(k)}(\lambda)\} \tag{13}$$

$$\lambda)=\mathfrak{I}_c^{-1}\{w_{n,k}\} \tag{14}$$

direct transform of $W^{(k)}(\lambda)$
$\lambda)=k$th derivative of a complex time function λ
discrete variable in the transform
half line corresponding to the complex time variable in the complex time plane
circle in the λ plane defining the domain of definition of the complex time function $W(k)(\lambda)$ and enclosing the singularities of the integrand of equation 11
complex time variable corresponding to the real variable t

ations 11 and 12 are used with a
of Taylor-Cauchy transforms to
nonlinear differential equations in
ner analogous to the use of Laplace
orms with linear constant parameter
differential equations. Details of this transform are given in reference 8 with adequate illustrations.

Consider now the conditions for the generalization of this transform to other partitions. Starting with equation 9 and again generalizing the real variable t to its analytic continuation in the complex λ plane, equation 15 is obtained:

$$x(\lambda)=\sum_{n=0}^{\infty} C_n Q_n(\lambda) \tag{15}$$

where

$x(\lambda)$ = a complex function derived from the real function $x(t)$ by analytic continuation just as $W(\lambda)$ was obtained from $x(t)$ in the case of the Taylor-Cauchy transform
C_n = corresponding transform mate of $x(\lambda)$

Multiplying equation 15 by $Q_{-m-1}(\lambda)$ and integrating around a contour C which encloses the appropriate singularities of the integrand gives interchanging orders of summation and integration.

$$\frac{1}{2\pi j}\int_C x(\lambda)Q_{-m-1}(\lambda)d\lambda=\sum_{n=0}^{\infty} C_n\frac{1}{2\pi j}\times\int_C Q_n(\lambda)Q_{-m-1}(\lambda)d\lambda \tag{16}$$

where the interchange is justified since the sum is uniformly convergent, under the conditions specified. To obtain a transform pair between $x(\lambda)$ and C_n, the sum on the right-hand side must be reduced to one term or, at most, a finite number of terms independent of λ. This last is easily obtained since the integration depends only on the residues of $Q_n(\lambda)\times Q_{-m-1}(\lambda)$ which are independent of λ. To obtain the former:

$$\frac{1}{2\pi j}\int_C Q_n(\lambda)Q_{-m-1}(\lambda)d\lambda=f(\delta_{n-m}) \tag{17}$$

must be satisfied where $f(\delta_{n-m})$ is some realizable function or operator f of the Kronecker delta function δ_{n-m}. In the Taylor-Cauchy transform, f is the idem-function and it occurs because $Q_n(\lambda)$ is a power function. Equation 17 is also satisfied when $Q_n(\lambda)$ has orthogonal properties. There are other properties of the moment function which will give rise to equation 17. Under these conditions equation 16 may be written as:

$$\sum_{n=0}^{\infty} C_n f(\delta_{n-m})=\frac{1}{2\pi j}\int_C x(\lambda)Q_{-m-1}(\lambda)d\lambda \tag{18}$$

When f is a linear operation with respect to δ_n:

$$\sum_{n=0}^{\infty} C_n f(\delta_{n-m})=f\left[\sum_{n=0}^{\infty} C_n\delta_{n-m}\right]=f(C_m) \tag{19}$$

whenever, in addition, the orders of summation and f operation are interchangeable. The general transform is therefore given by:

$$x(\lambda)=\sum_{n=0}^{\infty} C_n Q_n(\lambda) \tag{20}$$

$$f(C_n)=\frac{1}{2\pi j}\int_C x(\lambda)Q_{-n-1}(\lambda)d\lambda \tag{21}$$

The linear operation f on C_n, namely $f(C_n)$ and $x(\lambda)$, form a pair when properly treated and can be used in a manner similar to the Taylor-Cauchy transform for solving nonlinear differential equations describing physical nonlinear systems under the conditions specified in Appendix I. The functional form that the linear operator f takes is dependent upon the point of partition. The moment function $Q_n(t)$ is also dependent upon the point of partition. The summary for kth-order real system is given in Table I.

It is to be noted that the phase partitioning at a given derivative is taken to mean that the given derivative and all higher-order derivatives remain in the left member while all other terms are transposed to the right.

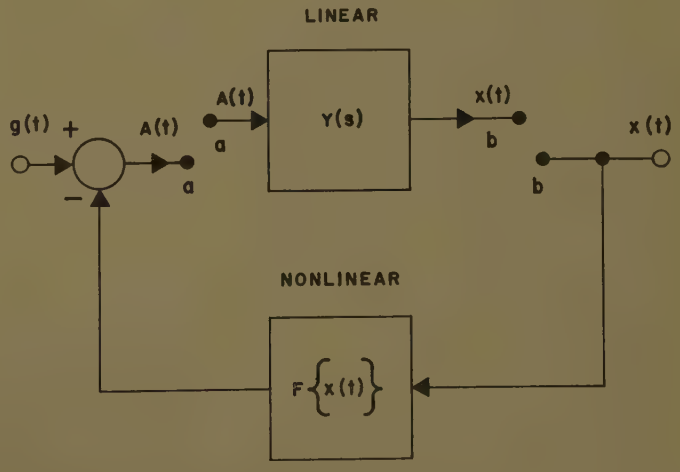


Fig. 3. The partitioned nonlinear system

EXAMPLE 1. PARTITIONING AN EQUATION

Given the nonlinear differential equation:

$$\frac{d^2x}{dt^2}+\frac{dx}{dt}+bx+cx^2=1 \tag{22}$$

1. Partitioning at the zero-th order derivative:

$$Q_n(t)=\frac{1}{\gamma_1-\gamma_2}\int_0^t \tau^n[e^{\gamma_1(t-\tau)}-e^{\gamma_2(t-\tau)}]d\tau \tag{23}$$

The integral of equation 23 can be evaluated by making use of:

$$\int_0^t \tau^n e^{-\gamma\tau}d\tau=(-1)^n\frac{d^n}{d\gamma n}\left[\frac{1-e^{-\gamma\tau}}{\gamma}\right]_0^t \tag{24}$$

where γ_1 and γ_2 are the roots of $s^2+as+b=0$.

2. Partitioning at the first derivative:

$$Q_n(t)=\frac{t^{n+1}}{n+1}-\int_0^t \tau^n e^{-a(t-\tau)}d\tau \tag{25}$$

3. Partitioning at the second derivative:

$$Q_n(t)=\frac{t^{n+2}}{(n+1)(n+2)} \tag{26}$$

EXAMPLE 2. APPLICATION OF TAYLOR-CAUCHY TRANSFORM

In this example apply the Taylor-Cauchy transform to case 3 of example 1. This gives:

$$w_n=\delta_n-\frac{aw_{n-1}}{n}-\frac{bw_{n-2}}{n(n-1)}-C\sum_{k=2}^{n-2}\frac{w_k w_{n-k-2}}{(k+1)(k+2)(n-k)(n-k-1)} \tag{27}$$

and, by looking for the corresponding pairs in a table of Taylor-Cauchy transforms,[8] one can obtain the time function $x(t)$ corresponding to the response of equation 22.

Random Processes in Physical Systems

Consider a physical system such as shown in Fig. 2 and suppose the forcing function $g(t)$ to be random with known statistical properties. The problem to be solved is to determine systematically the corresponding statistics of the response in terms of the statistics of the input. There are two theorems, proved elsewhere,[9,10] for what follows which are useful in the general synthesis of random processes.

TRANSFORM-ENSEMBLE THEOREM

Given a random process $g(t)$ with known statistical quantities and a deterministic operator L such that if $G(q)$ is the resulting transform-random process according to the relation

$$G(q)=L\{g(t)\} \tag{28}$$

where q is the transform-random variable corresponding to the orginal random t, then

$$\langle G(q)\rangle_G=L[\langle g(t)\rangle_g] \tag{29}$$

where $\langle\ \rangle_G$ is the ensemble average with respect to G and $\langle\ \rangle_g$ is the ensemble average with respect to g and commutative with respect to each other and the deterministic operator L, with the precaution that the subject of operation be the same for both operators.

The preceding theorem is useful for determining the first-order statistic of a system.[9]

EXAMPLE 3. LINEAR BROWNIAN MOTION SYSTEM

Let v be the velocity of each of a large number of similar but independently acting iron particles subjected to an oscillating magnetic force of magnitude a and a random fluctuating force $g(t)$ in a viscous medium giving rise to Doppler friction.[19] The differential equation governing this phenomenon is given by

$$\frac{dv}{dt}+\alpha v=g(t)+a\sin w_0 t \tag{30}$$

where

$\alpha=f/m$
f = the coefficient of friction
m = the mass of particle

Solution. In equation 29 let the operator L denote the Laplace operation. In a Brownian motion phenomenon the particles are all assumed to have started with the same initial velocity,[19] v_0. Thus:

$$V(s)=\frac{G(s)+a\dfrac{w_0}{s^2+w_0^2}+V_0}{s+\alpha} \tag{31}$$

Utilizing the inverse of equation 29 for this case

$$\langle v(t)\rangle_v=L^{-1}[\langle V(s)\rangle_V] \tag{32}$$

The result

$$\langle v(t)\rangle_v=L^{-1}\left[\frac{\langle G(s)\rangle_G}{s+\alpha}+a\frac{w_0}{(s+\alpha)(s^2+w_0^2)}+\frac{v_0}{s+\alpha}\right]$$

If $g(t)$ is stationary:

$$\langle G(s)\rangle_G=\frac{\langle g(t)\rangle_g}{s}$$

$$\langle v(t)\rangle_v=\langle g(t)\rangle_g[1-e^{-\alpha t}]+\frac{e^{-\alpha t}+\sin w_0 t-w_0\cos w_0 t}{(\alpha^2+w_0^2)w_0}+v_0e^{-\alpha t}$$

Higher-order statistics can easily b tained by appealing to the the given below. In order to discuss theorem, first denote by $\mathcal{C}^k\{\ \}$ the volution transform of the kth order the complex s plane this is:

$$\mathcal{C}\{F(s)\}=F(s)*F(s)$$

$$\mathcal{C}^2\{F(s)\}=F(s)*F(s)*F(s)\ldots \text{ etc.}$$

where * denotes the convolution in tion in the complex plane.

Higher-order statistic transform th Given a random process $g(t)$ a transform $G(q)$, the mean nth ens average of the given process is:

$$\langle g^n(t)\rangle_g=L^{-1}[\langle\mathcal{C}^{n-1}\{G(q)\}\rangle_G]$$

The proof of this theorem is ob by induction from the previous th with the aid of the convolution tran

EXAMPLE 4. SECOND STATISTIC SIMPLE STOCHASTIC PROCESS

Consider a simple stochastic p defined by the differential equatio initially at rest:

$$\frac{dx}{dt}=g(t)e^{-at};\ a>0$$

where $g(t)$ is an ergodic Gaussian ra process with zero mean and unit va The problem is to obtain the varia the response when the power-d spectrum is uniform over the spectrum.

Solution. The Laplace transfo equation 39 for $x(0)=0$ is:

$$X(s)=\frac{G(s+a)}{s}$$

Taking the convolution transfor

Table I

Partition Point	Moment Function
At kth derivative	power function
At kth$+(k-1)$ st derivatives	nonperiodic function of exponential type
At kth$+(k-2)$ nd derivatives	periodic function of exponential type
At all linear derivatives	combinations and linear combinations of power functions, and nonperiodic functions of exponential type

both sides of 40 and applying the foregoing theorem gives

$$\langle x^2(t)\rangle_x = L^{-1}[\langle \mathcal{C}X(s)\rangle_X]$$
$$= L^{-1}\left[\frac{\langle G(a)G(s+a)\rangle_G}{\blacksquare}\right] \quad (41)$$

To evaluate equation 41, use is made of the formula:[10]

$$\langle G(z+a)G(s-z+a)\rangle_G =$$
$$L_z L_s\{\langle g(t)g(t-\tau)\rangle_\varrho e^{-a\tau}e^{-2\alpha t}\} \quad (42)$$

where the complex variable s corresponds to t, a real variable, and z corresponds to τ, another real variable. Noting that the power-density spectrum is flat, i.e., white noise (equation 42) reduces to the simple result:

$$\langle G(z+a)G(s-z+a)\rangle_G = \frac{1}{s+2a} \quad (43)$$

Putting this into 41 gives:

$$\langle x^2(t)\rangle_x = L^{-1}\left[\frac{1}{s(s+2a)}\right] \quad (44)$$

or

$$\langle x^2(t)\rangle_x = \frac{1}{2a}[1-e^{-2at}] \quad (45)$$

Since the mean value of the process is zero, the variance σ^2 is:

$$\sigma^2 = \left[\frac{1}{2a}(1-e^{-2at})\right] \quad (46)$$

Relations between Wiener-Lee-Bose Theory and Wolf-Ku Theory

The Wiener theory of characterizing a nonlinear system consists, briefly, of obtaining a set of coefficients which are capable of physical measurement. These coefficients are obtained by subjecting the system under test to a white-noise probe. Because of the properties of white noise these coefficients determine uniquely the transmission characteristics of the given system. Since many systems with different physical configurations can exhibit the same transmission characteristics, the coefficients therefore do not uniquely determine the physical configuration of the system. The given system can, however, be found uniquely by a variational procedure which will be described.

Consider Fig. 4. This scheme was developed by Wiener, Lee, and Bose to test a nonlinear system so that information obtained from the response could be used to synthesize a new system having the same response to white noise as the original system. Since the power-density spectrum of white noise is flat over the infinite frequency interval, the system is effectively being exposed to every physical signal. It is well known that, in order to know the response of a nonlinear system to every input, it must be tested with every input.[2] This is unlike the behavior of a linear system that requires only an impulse as an input to determine its response to every input. This last occurs because of superposition. Hence, if the response to an arbitrary input is required, one would use the well-known convolution integral with the impulse response as a weighting function.

The Wiener theory fails to give back the original system configuration since the white-noise probe does not consider the effect of initial conditions. In linear systems the response is essentially independent of the initial boundary conditions. In nonlinear systems a change in initial conditions produces a change in the response. The most general test of nonlinear systems would be, then, to subject the input to every conceivable input and every conceivable initial value.

Fortunately, however, the use of white noise and parameter variation produces the same effect by generating a set of coefficients which not only uniquely defines the transmission, but also the configuration. Now consider the linear case.

Linear Case

Let $Y(s)$ be the transfer function of the physical system. Thus:

$$Z(D)x(t) = g(t) \quad (47)$$

describes the dynamic behavior of the linear system. Hence, by inversion:

$$x(t) = \int_{\blacksquare}^{\infty} g(t-\tau)y(\tau)d\tau \quad (48)$$

Similarly, from Fig. 4:

$$v_n(t) = \int_{\blacksquare}^{\infty} g(t-\gamma)h_n(\gamma)d\gamma \quad (49)$$

Multiplying equations 48 and 49 and averaging:

$$\overline{x(t)v_n(t)} =$$
$$\int_0^\infty \int_0^\infty \overline{g(t-\tau)g(t-\gamma)}y(\tau)h_n(\gamma)d\gamma d\gamma \quad (50)$$

Noting that white noise is a Gaussian process in this case, the outputs of the orthogonal filter are also Gaussian because

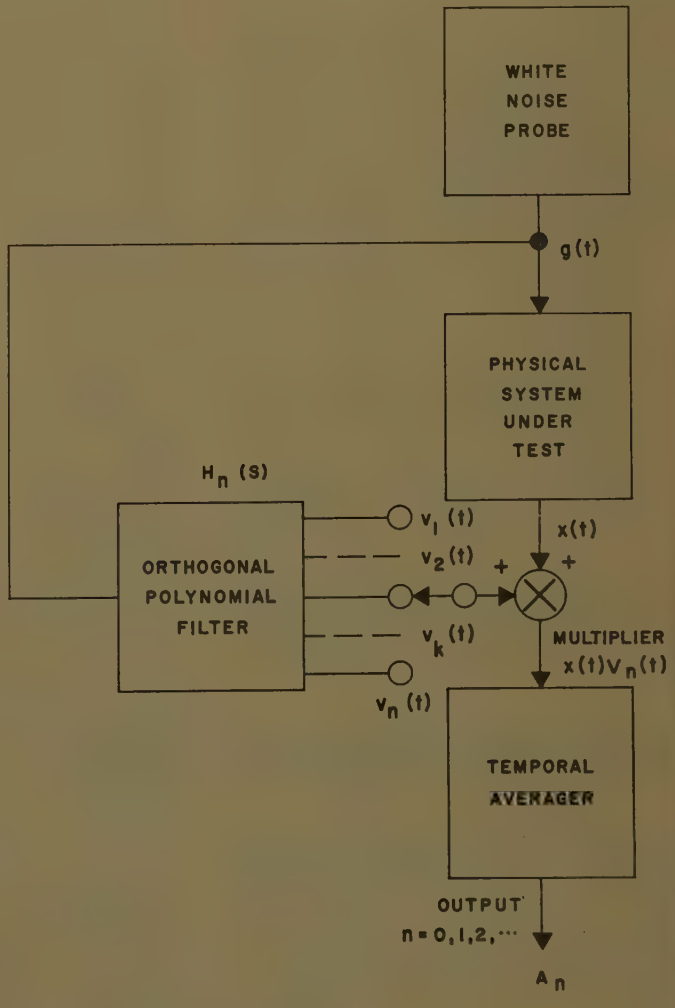


Fig. 4. The Wiener-Lee-Bose system

of linearity. Furthermore the outputs are statistically independent and all have the same variance. This obtains from the orthogonality and the nature of the network's nondissipation. Because of ergodicity, replace time averages with ensemble averages. Therefore equation 50 becomes

$$\langle x(t)v_n(t)\rangle_g = \int_0^\infty \int_0^\infty \langle g(t-\tau)\times g(t-\gamma)\rangle_g y(\tau)h_n(\gamma)d\tau d\gamma \quad (51)$$

The average inside the integral is recognized as the autocorrelation function. For white noise this is:

$$\langle g(-\tau)g(t-\gamma)\rangle_g = \delta(\tau-\gamma) \quad (52)$$

where $\delta(\tau-\gamma)$ denotes the Dirac delta function with power density.

Substituting equation 52 into equation 50 and simplifying gives:

$$\overline{x(t)v_n(t)} = \int_0^\infty y(\tau)h_n(\tau)d\tau \quad (53)$$

The left-hand side is obtained by the Wiener theory:

$$\overline{x(t)v_n(t)} = A_n \quad (54)$$

To display the dependence of y on its parameters:

$$y(\tau) = y(\tau;\ \alpha_1, \ldots, \alpha_k;\ \beta_1, \ldots, \beta_m) \quad (55)$$

where α_j and β_i are the time constants of the linear system. It is then evident that if $n=k+m$ measurements were made, the α_j and β_i could be uniquely determined from equation 56 by simultaneous solution of the set of algebraic equations which result.

$$A_n = \int_0^\infty y(\tau;\ \alpha_1, \ldots, \alpha_k;\ \beta_1, \ldots, \beta_m)\times h_n(\gamma)d\tau \quad (56)$$

To obtain the values of α_j and β_i in terms of the actual given components, use equation 57:

$$\delta A_n = \int_0^\infty \delta y(\tau;\ \alpha_1, \ldots, \alpha_k;\ \beta_1, \ldots, \beta_m)\times h_n(\tau)d\tau \quad (57)$$

where the variation δ is made with respect to the proper number of time constants according to the total number of components in the system.

Physical Interpretation of Integral Equations of Identification

Equations 56 and 64 are called the integral equations of identification since, as pointed out, the system parameters for linear and nonlinear systems can be determined from them by measurement of only the A_n statistics.

Examination of these equations shows that they are amenable to the solution of the system function $y(\tau)$, in the linear case, in terms of the orthogonal functions of $h_n(\tau)$ as linear combinations of the latter. That is, solving for $y(\tau)$ in equation 53 gives:

$$y(\tau) = \sum_{n=0}^\infty A_n h_n(\tau) \quad (53A)$$

subject only to the constraint given by

$$\int_0^\infty h_k(\tau)h_n(\tau)d\tau = \delta_{k-n} \quad (53B)$$

where, as before, δ_{k-n} is a Kronecker delta function.

It is therefore clear that equation 53(A) denotes a method of measuring the impulse response in terms of the impulse responses of the orthogonal filters by measuring only the statistics associated with the system under test, as depicted in Fig. 4, when the random source is stationary and white. Suppose now that $g(t)$ is an arbitrary random source so that its correlation function is

$$\phi(\tau-\gamma) = \langle g(t-\tau)g(t-\gamma)\rangle_g \quad (53C)$$

In terms of equation 53(C), equation 50 becomes

$$A_n = \int_0^\infty \int_0^\infty \phi(\tau-\gamma)y(\tau)h_n(\gamma)d\tau d\gamma \quad (53D)$$

If the system under test is fixed to, say, a straight-through connection, then:

$$y(\tau) = \delta(\tau) \quad (53E)$$

Thus:

$$A_n = \int_0^\infty \phi(\gamma)h_n(\gamma)d\gamma \quad (53F)$$

from which

$$\phi(\gamma) = \sum_{n=0}^\infty A_n h_n(\gamma) \quad (53G)$$

This gives a convenient method of measuring the autocorrelation function in a manner similar to system testing given in equation 53(A). Such a correlator has been tested experimentally and its results will be reported elsewhere.

Therefore, the integral equations of identification give rise to convenient methods of obtaining system attenuation and phase characteristics and correlation functions depending on whether the noise source is fixed or the system under test is fixed. Similar interpretation exists for the nonlinear case.

The Nonlinear Case

Consider equation 4 written as

$$x(t) = \int_0^\infty a(t-\tau)y(\tau)d\tau \quad (58)$$

where $a(t)$ is the modified forcing function. Let $a(t)$ be expanded into a series of complete orthogonal functions

$$a(t) = \sum_{k=0}^\infty a_k P_k(t) \quad (59)$$

where

$$\int_0^\infty P_k(t)P_m(t)dt = \delta_{k-m} \quad (60)$$

and

$$a_k = \int_0^\infty a(t)P_k(t)dt \quad (61)$$

Expansion 59 is justified since $P_k(t)$ is taken analytic for all k. Using equation 49 and multiplying and averaging:

$$\overline{x(t)v_n(t)} = \int_0^\infty \int_0^\infty \overline{a(t-\tau)g(t-\gamma)}\times y(\tau)h_m(\gamma)d\tau d\gamma \quad (62)$$

Now let

$$\overline{a(t-\tau)g(t-\gamma)} = \phi(\tau-\gamma) \quad (63)$$

so that

$$\overline{x(t)v_n(t)} = \int_0^\infty \int_0^\infty \phi(\tau-\gamma)y(\tau)h_n(\gamma)d\tau d\gamma \quad (64)$$

Solution of equation 64 requires a knowledge of the expansion of $\phi(\tau-\gamma)$. Consider one case patterned after the theory given in references 2 and 3. Let the right side of equation 63, ϕ, occupy the same place, with respect to a random process, as $a(t)$ occupies with respect to a deterministic process. Then, under certain conditions,[2-9] $\phi(\tau)$ can be expanded as follows:

$$\phi(\tau) = \sum_{n=0}^\infty C_n \tau^n \quad (65)$$

Noting that[21]

$$\phi(\tau-\gamma) = e^{-\gamma D}\phi(\tau) \quad (66)$$

where D is differential time operator $d/d\tau$ and substituting the right side of equation 65 into 66 for $\phi(\tau)$, results in:

$$\phi(\tau-\gamma) = e^{-\gamma D}\sum_{n=0}^\infty C_n \tau^n \quad (67)$$

Putting equation 67 into 64 gives:

$$\overline{x(t)v_n(t)} = \int_0^\infty \int_0^\infty \left(e^{-\gamma D}\sum_{n=0}^\infty C_n\tau^n\right)\times y(\tau)h_n(\gamma)d\tau d\gamma \quad (68)$$

Interchanging the order of summation and integration yields

$$\overline{x(t)v_n(t)} = \sum_{n=0}^\infty C_n \int_0^\infty \int_0^\infty (e^{-\gamma D}\tau^n)y(\tau)\times h_n(\gamma)d\tau d\gamma \quad (69)$$

is now convenient to define a new rameter, $M_n(\gamma)$, as:

$$_n(\gamma) = \int_0^\infty (e^{-\gamma D_\tau} \tau^n) y(\tau) d\tau \qquad (70)$$

$$_n(\gamma) = \int_0^\infty \tau^n y(\gamma+\tau) d\tau \qquad (70A)$$

This parameter represents the modified oments of $y(\tau)$, the impulse response of e linear part of the system for a random ocess, in the same manner that $Q_n(t)$ presents the moments of equation 9 for deterministic process. Using equation with 69 yields

$$\overline{)v_n(t)} = \sum_{n=0}^{\infty} C_n K_n \qquad (71)$$

ere K_n is defined by equation 72:

$$_n = \int_{\blacksquare}^{\infty} M(\gamma) h_n(\gamma) d\gamma \qquad (72)$$

e C_n's are obtained recursively follow- g a similar procedure to one given pre- ously.[2-4] Note that $x(t)$ can be ex- nded into a power series[9] such as:

$$) = \sum_{n=0}^{\infty} f_n t^n \qquad (73)$$

en the random process is a Brownian otion type. Then, noting that equa- n 59 can also be expanded into a power ries:

$$) = \sum_{n=0}^{\infty} b_n t^n \qquad (74)$$

ere b_n are the Taylor coefficients of the ght side of equation 59 namely:

$$= \frac{1}{n!} \frac{d^n}{dt^n} \left[\sum_{k=0}^{\infty} a_k P_k(t) \right]_{t=0} \qquad (75)$$

e last coefficients may also be obtained cursively. Rewriting equation 58 as

$$) = \int_0^\infty a(\tau) y(t-\tau) d\tau \qquad (76)$$

d using equation 74 yields:

$$) = \sum_{n=0}^{\infty} b_n Q_n(t) \qquad (77)$$

the partition is made at the highest- der derivative, $Q_n(t)$ is a power func- n.[2-3] In this case, recurrence relations e obtained relating the b_n to the a_n given equation 59. The f_n are automatically tained from equation 77. To relate e C_n to the b_n, appeal to equation 63. he latter yields the appropriate recur- nce relation to complete the solution. ie recurrence relations will not be dealt th in detail, however, a great deal of nplification can be effected in the form of the recurrence relations if $a(t)$ is assumed to be a power series at the start. In this case the Taylor-Cauchy transform can be applied and will now be considered with the detection of faults in complex physical systems.

General Synthesis Problem

In this section, a theory leading to a general synthesis procedure of both linear and nonlinear systems will be considered, the problem being that of uniqueness.

In the problem of analysis the solution is always unique for linear systems and for a certain class of nonlinear systems which are physically realizable.[2-3] In the problem of synthesis the solution is not generally unique. The reason for this lack of uniqueness seems to be related to a corresponding lack of information concerning such things as initial conditions, forcing functions, distribution of currents flowing in the interior of the structure and the number of components comprising the configuration. Generally, in linear synthesis procedures, the forcing function is tacitly given when the procedure involves poles and zeros in the complex frequency plane, otherwise none of the other information is usually given. This allows for an ambiguity in construction of the configuration giving rise to as many as an infinite variety of solutions.

In the previous section it was pointed out that the Wiener theory by itself led to a synthesis procedure which was not unique. The Wiener structure combined Laguerre and Hermite networks with multipliers and adders to produce a response to white noise which was identical to the system under test. By combining the Wiener-Lee-Bose theory with the Wolf-Ku theory a set of measured values are related to the transmission coefficients of the structure which are in turn related to the time constants of the structure. However, this knowledge does not give a unique structure of the system under test since once again many systems can have equivalent transmission with such time constants. To define the nature of these time constants in order to define uniquely the structure, the system should be tested in the following way. Expose the system to a white-noise probe as a forcing function and, simultaneously, vary the boundary conditions according to another white-noise probe generated independently; see Fig. 5. The unique synthesis is obtained by a statistical analysis of the output. This might be achieved by subjecting the output to an analysis similar to the Wiener scheme with the analysis repeated with orthogonal components obtained from both white-noise probes. This problem has not yet been satisfactorily solved. The main difficulty stems from the introduction of the second noise probe.

As remarked earlier, if the structure configuration is given a priori, the unique synthesis problem is reduced to a procedure involving two sets of measurements of the Wiener type; the second is obtained by repeating the first after a variation has been performed on the system in a specified manner.

Synthesis to Special Waveforms-Pole-Zero Method

In the previous section the synthesis of networks and systems under the influence of a white-noise probe was discussed. Now, a synthesis procedure is indicated for nonlinear systems given a specified waveform as an input, boundary values, a distribution of singularities in the complex plane, and the transmission characteristics of the linear part. The procedure is based on a theory outlined for stable systems in reference 2 and continued in reference 7, however, this procedure has a much simpler concept.

Let $Y(s)$ be the transfer function of the linear part and let $\{s_n\}$ be a set of singularities which are generally supposed to be finite in number; then, given a specified forcing function $g(t)$, what is the nature of the nonlinear part F to satisfy these conditions?

The synthesis procedure under these conditions is:

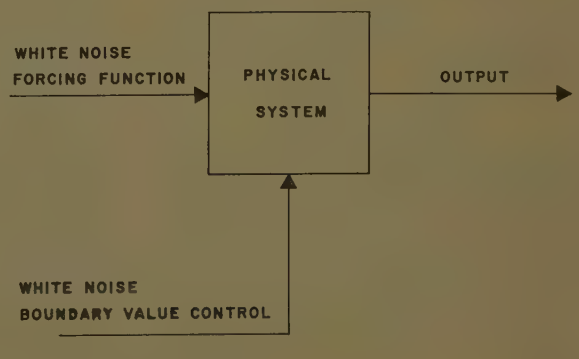


Fig. 5. A procedure for uniquely determining the synthesis of a given structure

1. From the singularities and initial conditions determine the coefficients C_n in closed form using reference 2.

2. Utilizing the C_n coefficients, determine the response of the system using the moment theorem.

3. The nonlinear part of the system is then obtained using the canonical equation 1.

4. The canonical system is then obtained easily by reference to Fig. 2.

EXAMPLE

Let the linear plant of a system be given as

$$Y(s)=\frac{1}{s+2} \tag{78}$$

If the system is forced by

$$g(t)=e^{-2t} \tag{79}$$

find the coefficients of the nonlinear terms so that the response has a pole at $s=-1$ and the nonlinear part has no derivatives and does not exceed degree 2.

The solution can be effected by using the recurrence equations of references 2–4, or it may be solved directly, in which case the solution is almost trivial. For instance, from

$$Z(S)=\frac{1}{Y(S)}=S+2 \tag{80}$$

the canonical form is obtained:

$$\frac{dx}{dt}=2x+F(x)=g(t) \tag{81}$$

or

$$F(x)=g(t)-\dot{x}-2x \tag{82}$$

Noting

$$F(x)=\sum_{k=0}^{\infty} a_k x^k$$

then

$$\sum_{k=0}^{2} a_k x^k=g(t)-\dot{x}-2x \tag{83}$$

From the problem:

$$X(s)=\frac{1}{s+1} \tag{84}$$

from which

$$x(t)=e^{-t} \tag{85}$$

so that

$$a_0+a_1e^{-t}+a_2e^{-2t}=e^{-2t}-e^{-t} \tag{86}$$

Since equation 86 is an identity, it is clear that:

$$a_0=0 \tag{87}$$

$$a_1=-1 \tag{88}$$

$$a_2=1 \tag{89}$$

Thus the nonlinear component satisfying the problem requirements is:

$$F(x)=-x+x^2 \tag{90}$$

Conclusions

This paper has presented some aspects of some of the recent advances in the analysis and synthesis of nonlinear systems. The concepts involved in the Wiener theory and those developed by the author have been reviewed. It is evident that the two theories have different approaches looking toward common goals. On one hand, the Wiener theory considers the synthesis problem in a statistical form, while on the other hand the author's theory considers the concept of partition which is a deterministic process. However, the two theories may be combined to produce a new theory for determining the structural configuration internal to a given system. This theory therefore finds application in the fault detection and diagnosis problem in complex systems. In addition to its practical value the new theory relates coefficients capable of being measured to coefficients capable of being calculated, which gives a means of mechanizing new measuring systems with high accuracy.

By interpreting the solutions of the integral equations of identification, one can devise highly accurate measuring systems to determine attenuation and phase characteristics for linear systems, and correlation functions of stationary random functions. This leads to the concept of "almost instantaneous" and "instantaneous" correlators.

A complete synthesis theory can be developed using the canonical system as a basis. From here it is easy to show that the response of this system is unaltered if the linear and nonlinear boxes are interchanged providing each operator is also inverted. It is possible to show that by making the appropriate partitions almost any desired configuration can be achieved in both single- and multiple-loop feedback systems.

It is interesting to note that the absolute test for stability of a nonlinear system consists of utilizing a double white-noise probe (Fig. 5) and examining the position of the resulting singularities in the complex frequency plane as described in reference 7.

Some very important problems are still left unsolved; for instance, the solution of the recurrence equations in closed form, development of other transforms like the Taylor-Cauchy to give elementary and higher transcendental types of solutions for nonlinear systems, and using more efficiently Kronecker's theorem (developed in reference 2) as a tool in the synthesis of nonlinear systems.

It is possible to extend the synthesis problem, under suitable restrictions, to the design of adaptive control systems.

Appendix I. Restrictions on System

1. The linear operator $Z(D)$ has an impulsive response $y(t)$ of exponential type and order 1.

2. The forcing function $g(t)$ is a function of exponential type and order 1.

3. The nonlinear function $F(x, \dot{x}, \ldots)$ is single-valued, continuous, and satisfies the following conditions.

Given a pair of positive constants, M and a, and two continuous functions, $v(t)$ and $u(t)$, which are asymptotically like e^{-at}, then F must satisfy the condition that

$$|F\{v(t)\}-F\{u(t)\}|<Me^{-at}|v(t)-u(t)|$$

for all $t>0$.

Appendix II. Partitioning to Obtain Power-Series Solutions

In the partition theory of solving ordinary nonlinear differential equations of certain class, all the nonlinear terms are transposed to the right side. The solution is a linear combination of the moments of the folded impulse response of the linear terms in the open interval $(0, t)$. It is noted that if all but the highest-order derivative is transposed to the right member, power-series solution results because the resulting moment function is a power function. It is possible to obtain power-series solutions by transposing all terms except any one linear term to the right side. However, in the case of linear terms of order less than the highest-order derivative, it is necessary to reorder the indexes prior to obtaining the recurrence equation. If prior reordering of indexes is not according to the order of the term remaining in the partition, the resulting recurrence relation will be backward.

The partitioning of an equation at the highest derivative has the benefit of giving the desired coefficients automatically as a function of the lower-order ones without prior reordering of indexes.

References

1. THEORY OF NONLINEAR CONTROL, Y. H. Ku "Proceedings, First International Symposium on Automatic Control," Moscow, Aug. 1960, Butterworth's Scientific Publications, London, England, 1961; also *Journal*, Franklin Institute, Philadelphia, Pa., vol. 271, Feb. 1961, pp. 108–44.

2. A MATHEMATICAL THEORY FOR THE ANALYSIS OF A CLASS OF NONLINEAR SYSTEMS, Alfred A. Wolf, *Doctoral Dissertation*, The University of Pennsylvania, Philadelphia, Pa., June 1958.

3. RECURRENCE RELATIONS IN THE SOLUTION OF A CERTAIN CLASS OF NONLINEAR SYSTEMS, Alfred A. Wolf. *AIEE Transactions*, pt I (*Communication and Electronics*), vol. 78, 1959 (Jan. 1960 section), pp. 830–34.

GENERALIZED RECURRENCE RELATIONS IN THE ALYSIS OF NONLINEAR SYSTEMS, Alfred A. Wolf. id., vol. 80, Sept. 1961, pp. 383–87.

ANALYSIS OF TRANSCENDENTAL NONLINEAR STEMS, Alfred A. Wolf. *Ibid.*, vol. 79, Nov. 1960, . 449–51.

ON THE SIGNIFICANCE OF TRANSIENTS AND EADY STATE BEHAVIOR IN NONLINEAR SYSTEMS, A. Wolf. *Proceedings*, Institute of Radio En- neers, New York, N. Y., vol. 47, no. 10, Oct. 1959, . 1785–86.

A STABILITY CRITERION FOR NONLINEAR SYS- MS, Y. H. Ku, A. A. Wolf. *AIEE Transactions*, . II (*Applications and Industry*), vol. 78, July 59, pp. 144–48.

TAYLOR-CAUCHY TRANSFORMS FOR ANALYSIS A CLASS OF NONLINEAR SYSTEMS, Y. H. Ku, A. Wolf, J. H. Dietz. *National Convention ecord*, Institute of Radio Engineers, New York, Y., pt. 2 (*Circuit Theory*), 1959, pp. 49–61; also *roceedings*, Institute of Radio Engineers, vol. 48, . 5, May 1960, pp. 912–22.

TRANSFORM-ENSEMBLE METHOD FOR THE ANALYSIS OF LINEAR AND NONLINEAR SYSTEMS WITH RANDOM INPUTS, Y. H. Ku, A. A. Wolf. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 15, 1959, pp. 441–45.

10. ON POLES AND ZEROS OF A RANDOM PROCESS IN LINEAR AND NONLINEAR SYSTEMS, A. A. Wolf. *Ibid.*, vol. 16, 1960, pp. 268–78.

11. A THEORY OF NONLINEAR SYSTEMS, A. G. Bose. *Technical Report no. 309*, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., May 15, 1960.

12. NONLINEAR PROBLEMS IN RANDOM THEORY (book), N. Wiener. John Wiley & Sons, Inc., New York, N. Y., 1958.

13. CYBERNETICS (book), N. Wiener. John Wiley & Sons, Inc., 1948.

14. APPLICATION OF STATISTICAL METHODS TO COMMUNICATION PROBLEMS, Y. W. Lee. *Technical Report no. 181*, Research Laboratory of Electronics, Massachusetts Institute of Technology, Sept. 1, 1950.

15. THEORY OF NONLINER TRANSDUCERS, H. E. Singleton. *Technical Report no. 160*, Research Laboratory of Electronics, Massachusetts Institute of Technology, Aug. 12, 1950.

16. THEORIE AND ANWENDUNG DER LAPLACE TRANSFORMATION (book), G. Doetsche. Dover Publications, Inc., New York, N. Y., 1943.

17. ENTIRE FUNCTIONS (book), R. P. Boas. Academic Press, New York, N. Y., 1954.

18. ON A SYSTEMATIC APPROXIMATION TO THE PARTITION METHOD FOR ANALYSIS OF A CLASS OF NONLINEAR SYSTEMS, Y. H. Ku, A. A. Wolf, J. H. Dietz. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 79, July 1960, pp. 183–91.

19. ON THE THEORY OF THE BROWNIAN MOTION, G. E. Uhlenbeck, L. S. Ornstein. *Physical Review*, New York, N. Y., Sept. 1, 1930, pp. 823–41.

20. THE FOURIER INTEGRAL AND CERTAIN OF ITS APPLICATIONS (book), N. Wiener. Dover Publications, Inc., 1933.

21. THE CONVOLUTION TRANSFORM (book), D. V. Widder, I. I. Hirschman. Princeton University Press, Princeton, N. J., 1955.

Discussion

. H. Ku (Moore School of Electrical Engineering, University of Pennsylvania, hiladelphia, Pa.): In connection with the artition theory discussed in the beginning Dr. Wolf's paper, I did suggest the name "auxiliary forcing function" for $A(t)$ used equations 4 and 5 and shown in the block agram, Fig. 2, of the paper. I wish to oint out that in the nonlinear differential quation 1, transposition is allowable for a ass of problems where the solution is nique and analytic, although superposition ould certainly fail to apply. In a recent aper[1] I have discussed this point in connecon with the feedback system block diaram. Let the nonlinear differential equaon be

$$'+k_1c'+k_2c=r-N(c',c)=r-b=e \qquad (91)$$

here c denotes the output signal, c' and denote the first and second derivatives of with respect to time, and $N(c', c)$ denotes nonlinear function of c' and c. The symols r, b, and e correspond to the input, feedack signal, and error or actuating signal spectively. Note that the actuating signal t) corresponds to a particular choice of the uxiliary forcing function $A(t)$ mentioned Dr. Wolf's paper.

We can rewrite equation 91 in two differnt ways:

$$'+k_1c'=r-k_2c-N(c',c)=r-b_1=e_1 \qquad (92)$$

$$'=r-k_1c'-k_2c-N(c',c)=r-b_2=e_2 \qquad (93)$$

hus, e_1 and e_2 correspond to $A_1(t)$ and $A_2(t)$, ne two other modified auxiliary functions, spectively. Suitable block diagrams can e drawn corresponding to equations 2 and

So the method is not limited to partioning at the highest derivative alone, hough the highest derivative should be inuded on the left-hand side of equations 91–3. Note that the original nonlinear differntial equation would be

$$'+k_1c'+k_2c+N(c',c)=r \qquad (94)$$

hich may represent an open-loop system. The advantage of the Taylor-Cauchy ansform method is in its possibility of any applications. For instance, it can be sed with the Transform-Ensemble method mentioned in reference 9 of the paper. It can be readily used for the solution of varying-parameter systems.[2] The writer has great hope in Dr. Wolf's extension of the above-mentioned approach to the synthesis of nonlinear systems. Maybe it is in the field of synthesis rather than the field of analysis that the new method will be most useful.

REFERENCES

1. ON NONLINEAR NETWORKS WITH RANDOM INPUTS, Y. H. Ku. *Transactions*, Institute of Radio Engineers, New York, N. Y., vol. CT-7, no. 4, Dec. 1960, pp. 479–90.

2. TAYLOR-CAUCHY TRANSFORMS FOR ANALYSIS OF VARYING-PARAMETER SYSTEMS, Y. H. Ku. *Proceedings*, Institute of Radio Engineers, vol. 49, no. 6, June 1961, pp. 1096–97.

R. L. Cosgriff (Ohio State University, Columbus, Ohio): The author is to be congratulated for focusing attention upon series solutions of nonlinear differential equations. This discussion is concerned with the real time series solutions and the comparison of this technique with the Taylor-Cauchy transform method.

Nonlinear differential equations with deterministic driving functions can be solved as accurately as desired providing one is willing to expend the energy and a solution exists. Many methods can be employed, and a judicious choice of method for a given problem can greatly reduce the labor involved. One of the classical methods for the solution of linear differential equations is the method of Frobenius, that gives the solution in terms of a Taylor series. For a complete discussion see reference 1. This method when extended to the nonlinear[2] case yields results identical to those obtained by the Taylor-Cauchy transforms. Disregarding terminology, the basic steps of this transform technique and those involved in the method of Frobenius are identical. In both cases a Taylor expansion of all variables involved in the differential equation must exist. (The collective coefficients of the series for each of the various variables is the transform of that variable.) Equating all coefficients, known and unknown, of like powers of t arising after the substitution of time series into the differential equation corresponds to equating the transforms of the variables involved. Determining the transform of the unknown variable corresponds to determining the coefficients of the time series of this unknown variable.

Consider the author's example

$$\frac{d^2x}{dt^2}+a\frac{dx}{dt}+bx+cx^2-1=0$$

$$x=\dot{x}=0 \text{ at } t=0$$

Let

$$x=\sum_{n=2}^{\infty} w_n t^n$$

$$w_o=w_1=0$$

the series for x^2 becomes

$$x^2=\sum_{n=2}^{\infty}\sum_{m=2}^{n-2} w_m w_{n-m} t^n$$

Upon substitution of the assumed series one has

$$\sum\left[n(n-1)w_n+a(n-1)w_{n-1}+bw_{n-2}+c\sum_{m=2}^{n-4} w_m w_{n-m-2}\right]t^{n-2}-1=0$$

Thus, the coefficients for the unknown x (the transform of x) are given by

$$w_n=-\frac{aw_{n-1}}{n}-b\frac{w_{n-2}}{n(n-1)}-c\sum_{m=2}^{n-4}\frac{w_m w_{n-m-2}}{n(n-1)}$$

for $n>2$

The above recursive relationship gives the unknown coefficients (the transform) of x and is identical to equation 27 once the w's of equation 27 are properly defined or the expression is corrected. For $a=b=c=1$, $w_2=1/2$, $w_3=-1/16$, $w_4=0$, $w_5=1/120$, $w_6=-1/720-1/120=-7/720$, therefore

$$x=t/2^2-1/6t^3+1/120t^5-7/720t^6+\ldots$$

Notice the two terms of w_6. The first would occur if $c=0$ (the linear differential equation) and the second is due to the nonlinear term. Observe the large magnitude due to the nonlinear term. For $n>6$ the nonlinear terms predominately determined the values of the coefficients of the differential equation. Thus, it can be concluded that the rate of convergence must be considered,

and one can easily demonstrate that this is far more of a limitation in the solution of nonlinear problems than it is in the case of linear problems.

Returning to the author's development it is apparent that the special techniques such as partition theory, moment theory, and the Taylor-Cauchy transform can be bypassed by using the method of Frobenius.

The author considers desirable requirements for other transforms equation 16 through 21. These conditions describe functional groups convenient for the solution of differential equations. In the real domain desirable requirements for functional groups can be described; for example,

$$\frac{df_n}{dt}=\sum_{M_1}^{M_2} k_m f_m \qquad M_1 \leq M_2 \leq n$$

$$f_m f_n=\sum_{\alpha}^{\beta} K_q f_q$$

with the requirement $m \leq n \leq \alpha \leq \beta$ where k's and K's are constants. If these conditions are met and the driving function, y, can be expanded in terms of the summation

$$y=\sum c_{dn} f_n(t)$$

the solution for x will take the form

$$x=\sum c_{sn} f_n(t)$$

and the recursive relationships for the c_{sn} will exist. Simple functions are

$$f_n=t^n$$

$$f_n=e^{-\alpha n t} \qquad (\alpha \text{ is a constant})$$

$$f_n=t^n e^{-\alpha t}$$

I wonder whether these real time requirements are the same as the requirements based upon transform theory.

The importance of Wolf's work with transform methods is that it introduces a new facet concerning the solution of nonlinear differential equations. As each facet is exploited new insights and techniques develop which expand our knowledge of nonlinear systems.

References

1. Ordinary Differential Equations, (book), **E. L. Ince.** Dover Publications, Inc., New York, N. Y., pp. 396–403.

2. Nonlinear Control Systems (book), **R. L. Cosgriff.** McGraw-Hill Book Company, Inc., New York, N. Y., 1958, pp. 15–16.

Louis F. Kazda and **A. Y. Bilal** (University of Michigan, Ann Arbor, Mich.): Dr. Wolf has called the attention of control system engineers to the partitioning method and its application to nonlinear differential equations. It is a systematic way of solving nonlinear differential equations, and lends itself to digital computation. Probably the greatest contribution of the method is that it permits an insight into the synthesis of nonlinear systems.

A review of the author's published work, however, would reveal that certain questions have been left unanswered, and which, the novice, trying to utilize the presented material, would naturally like to have summarized. These questions are:

1. What insight can be gained about the stability of a system for a general class of forcing functions utilizing the partitioning method of analysis?

2. Under what conditions would partitioning at points other than the highest derivative be desirable?

3. Consider for example the nonlinear differential equation

$$\ddot{x}+\omega_0 x+x^3=0$$

which is the equation that is obtained as a result of introducing into the linear second-order differential equation $\ddot{x}+\omega_0 x=0$ the nonlinear term x^3. It is not clear to us how the partitioning method presented will aid in telling us how the amplitude and frequency of oscillations vary with time.

Alfred A. Wolf: I am most grateful to Dr. Ku for his illuminating and encouraging discussion. In general we appear to be in agreement on all the points discussed. In particular, Dr. Ku's reference to the synthesis of nonlinear systems is of particular interest. In my dissertation,[1] a stability theory was given. There are several interesting consequences of this theory, one of which was reported upon previously.[1,2] I would now like to discuss briefly another consequence that leads to a synthesis procedure.

Suppose the Laplace transform of $x(t)$, the solution or response of the system, after suitable change of variable as discussed in reference 1 is obtained by the Taylor-Cauchy transform[3] or some other method[1] and given by

$$X(w)=\sum_{n=0}^{\infty} h_n w^n \tag{95}$$

where w is a complex variable corresponding to t and functionally related to the complex variable, s, of the Laplace transform and h_n denotes the coefficients of w^n.

Under certain conditions $X(w)$ will be meromorphic. A simple test for determining this is by applying Kronecker's theorem.[1,4] Then $X(w)$ is expressible as the rational fraction,

$$X(w)=\frac{N(w)}{P(w)} \tag{96}$$

where

$$N(w)=\sum_{r=0}^{m} a_r w^r \tag{97}$$

and

$$P(w)=\sum_{r=0}^{n} b_r w^r \tag{98}$$

then

$$a_r=\sum_{k=0}^{r} h_k b_{r-k} \tag{99}$$

such that

$$a_r \begin{cases} =0 \text{ when } r>m \\ \neq 0 \text{ when } r \leq m \end{cases} \tag{100}$$

The upper condition of equation 100 determines the b_r's while the lower condition determines the a_r's once the b_r's are known. It is evident therefore that it is possible to calculate the zeros and poles of $X(W)$. Thus, the stability is determined.

Example

Let us determine the coefficients a_k and b_k for the following problem in terms of the coefficients h_n.

$$\sum_{n=0}^{\infty} h_n w^n=\frac{a_0+a_1 w}{b_0+b_1 w} \tag{101}$$

$$=\frac{a_0}{b_0}\frac{1+\frac{a_1}{a_0}w}{1+\frac{b_1}{b_0}w}$$

Therefore

$$a_0=h_0 b_0 \tag{102}$$

$$a_1=h_0 b_1+h_1 b_0 \tag{103}$$

$$0=h_1 b_1+h_2 b_0 \tag{104}$$

$$0=h_2 b_1+h_3 b_0 \tag{105}$$

For nontrivial solution a condition on the h's:

$$\begin{vmatrix} h_1 & h_2 \\ h_2 & h_3 \end{vmatrix}=0 \tag{106}$$

Hence

$$\frac{b_1}{b_0}=\frac{h_2}{h_1} \tag{107}$$

$$\frac{a_0}{a_1}=h_0 \tag{108}$$

$$\frac{a_1}{a_0}\frac{h_0 h_1}{h_0 h_2+h_1^2} \tag{109}$$

The h_n's are then determined in terms of the a_n's. Turning the problem around we have the synthesis problem; that is, given the a_n's, which specify the characteristics of a certain system, the h_n's are calculated. The dynamics of the system would be specified by a nonlinear differential equation in which at least some of the coefficients of the differential equation are unknown. The h_n's are then functions of these unknown coefficients. Since the a_n's are specified, the h_n's are determined, hence the unknown coefficients describing the system dynamics are calculated. This is one kind of synthesis procedure leading to systems with a specified degree of stability and the parameters of the system are determined in such a way as to give this desired stability.

I would like to thank Dr. Cosgriff both for his comments and for affording me the opportunity of replying to the points raised.

The two main points which Dr. Cosgriff raised deal with the method of Frobenius and its relation to the method of the paper and the generalization of expanding the auxiliary forcing function.

The method of Frobenius, which has been discussed in some detail in reference 5, is a special case of the Partition Theory. However, the Frobenius method only gives rise to power series solution, but the Partition Theory gives rise to a general class of solution forms depending on the point of parti-

. As pointed out in the paper, these
ıs can be Power series, Dirichlet series,
gonometic series, Orthogonal series,
ng others. The form of solution de-
ds on the partition point and the form of
ansion of the auxiliary function, which
vide flexible control on the nature of
solution's form. In the Frobenius
hod the solution is assumed to be a power
es; in the partition scheme the auxiliary
ction is expanded into an analytic series,
necessarily a power series. If the auxil-
function is expanded into a power series,
solution is only a power series if the
ıt of partition is at the highest derivative.
vill also be a power series if only one term
ains after partition and reordering of
ices. If the partition does not allow the
ıest order derivative to remain on the
member then care must be taken so that
resulting recurrence equation will not be
kward. In this sense, a partition at the
ɔ-th derivative, with all other linear
ns transposed, is equivalent to the
thod of Frobenius. In the paper, the
ct of changing the partition point was
strated. This illustration showed how
solution form changed from a power
es to a Dirichlet series.
We now turn to the second point of Cos-
F's discussion dealing with the con-
uction of analytic auxiliary functions
er than power series. This point is dis-
sed in great detail in reference 1. In-
ad of going into such detail here an ex-
ple will be given.

AMPLE

In virtue of the theory in reference 1,
following expansion of A(t) is justified.

$$t)=\sum_{n=0}^{\infty} c_n e^{-\lambda_n t} \tag{110}$$

our previous examples $A(t)$ was ex-
nded into a power series. The recurrence
ations resulting from this kind of expan-
n are simple and elegant. The next
mplest expansion of $A(t)$ is given by equa-
n 110 and the simplest recurrence
ations among these results when λ_n is
osen as

$$=an \tag{111}$$

uation 1 now has the solution

$$)=\sum_{n=0}^{\infty} c_n P_n(t) \tag{112}$$

ere $P_n(t)$ are the exponential moments
the folded impulsive response of the
ar part given by

$$t)=\int_0^t e^{-nar} y(t-\tau)\,d\tau \tag{113}$$

is evident that on expanding $e^{-na\tau}$ into
power series a relation exists between
exponential moments and the moments
$(t-\tau)$, $Q_n(t)$.
By choosing λ_n according to equation 111,
have the special case of the Dirichlet
es, known as the Power Dirichlet series.
s terminology is adapted since equation
ices to a power series in Z when $Z=$
^{t} and λ_n is given by equation 111. Con-
r, for example, a system described by

$$\frac{dx}{dt}+fx+bx^2=g \tag{114}$$

where $f, b,$ and g are constants, subject to the initial conditions

$$x(0)=0 \tag{115}$$

Equation 114 may be partitioned in two ways. The first might be at the highest derivative or a second choice might include, in addition to the derivative, the function itself. Consider the former partition first:

$$\frac{dx}{dt}=g-fx-bx^2 \tag{116}$$

Thus,

$$\frac{dx}{dt}=\sum_{n}^{\infty} c_n e^{-nat} \tag{117}$$

performing the indicated operations and equating like powers of $Z=e^{-at}$ the following recurrence relation is obtained.

For $n>0$:

$$C_n=(g-fk-bk^2)c_n^{(0)}+\frac{fc_n}{na}-bc_{n,0}{}^{(2)}+2bk\frac{c_n}{na} \tag{118}$$

where k is a constant of integration.

For $n=0$ we are enabled to determine k; that is, letting $n=0$ we obtain

$$bk^2+fk-g=0 \tag{119}$$

or

$$k=\frac{-f\pm\sqrt{f^2+4bg}}{2b} \tag{120}$$

For $n=1$:

$$c_1=f\frac{c_1}{a}-bc_{1,0}{}^{(2)}+2bk\frac{c_1}{a} \tag{121}$$

or

$$c_1\left(\frac{f}{a}+\frac{2bk}{a}\right)=0 \tag{122}$$

The second factor is not zero; hence,

$$c_1=0 \tag{123}$$

For $n=2$ it is possible to determine the characteristic exponent, a. After simplification

$$c_2\left[1-\frac{f}{2a}+\frac{bc_1{}^2}{a^2}-\frac{2bk}{2a}\right]=0 \tag{124}$$

Either or both factors may be zero. We shall assume that the second factor is zero to determine a. Noting $c_1=0$ from equation 123 and substituting the value of k given by equation 120 yields the results

$$a=\pm\tfrac{1}{2}\sqrt{f^2+4bg} \tag{125}$$

This result is known to be correct since it can be checked against the solution as obtained by separating the variables. The characteristic exponent in this case has two values. Both must be used in equation 126 giving rise to Dirichet series

$$x=\sum_{n=1}^{\infty}\frac{c_n}{-na}e^{-nat}+k \tag{126}$$

If the characteristic exponent had k values then k series would be needed in addition to the constants of integration which may arise. This procedure is similar to that used in linear differential equations. The remainder of the solution follows a similar procedure to that described elsewhere.

If the second partition scheme is used instead, a different form of $x(t)$ is obtained; this is easily checked.

The third case proposed by Cosgriff, namely expanding

$$A(t)=\sum_{n=0}^{\infty} c_n t^n e^{-at}$$

evidently follows a similar procedure. Actually a more interesting case for investigation is to expand $A(t)$ according to the relation

$$A(t)=\sum_{n=0}^{\infty} c_n t^n e^{-nat} \tag{127}$$

Regarding the question of generating transforms for these cases like the Taylor-Cauchy transform, the answer is clearly that this is indeed possible and has been done. These results may be reported upon later.

The questions raised by Dr. Kazda and Dr. Bilal are much appreciated and greatly valued. The three points raised by them are particularly pertinent to the theory of partition and I shall take up certain aspects of these points not covered before.

Let us consider the first point relative to the stability of nonlinear systems. In reference 1 it was shown that a certain class of linear systems existed (not necessarily physically realizable) such that the response, $x_n(t)$ for $n=0, 1, 2, \ldots$, of each of these systems formed a point set whose limit as $n\to\infty$ is the response of the nonlinear differential equation given in the paper as equation 1. The members of the set $x_n(t)$ are generated as follows:

$$x_0(t)=\int_0^t g(\tau)y(t-\tau)\,d\tau \tag{128}$$

and

$$x_n(t)=x_0(t)-\int_0^t F\{x_{n-1}(\tau)\}y(t-\tau)\,d\tau$$

for $n\geqslant 1$ (129)

The limit referred to,

$$x(t)=\lim_{n\to\infty} x_n(t) \tag{130}$$

is rigorously proved in references 1 and 6 as a uniformly convergent procedure. Because of uniformity, the following theorem is proved.

Theorem 1. Suppose that

$$\lim_{t\to\infty} x_n(t)=0 \tag{131}$$

then there exists an $N>0$ such that for each $n\geqslant N$ the limit, equation 131, governs the behavior of $x(t)$ so that

$$\lim_{t\to\infty} x(t)=0 \tag{132}$$

What this theorem implies is that there is, from a certain member on, in the set of responses $\{x_n(t)\}$ a linear member whose stability governs the stability of the nonlinear system. While theoretically this is an important result, the solution of $x_n(t)$ for a given $n \geqslant N$ is cumbersome and its stability is not immediately discernible. To ameliorate this situation one makes use of the following theorem.

Theorem 2. The members of the point set $\{x_n(t)\}$, under the conditions specified in the paper, are each of exponential type.

This implies that each member has a Laplace transform, and moreover, the singularities for a lumped parameter system are poles for the $\mathcal{L}\{x_n(t)\} = X_n(s)$ for all finite n. In the limit the function $x(t)$ is not necessarily of exponential type so it does not necessarily have a Laplace transform.

Utilizing the partition method the series solution of $x(t)$ can be obtained. By proper partition, the power series solution, or any other series solution is obtainable. Even though the expansion of $x(t)$ does not have a Laplace transform directly it can be obtained as near as desired by a limiting procedure on the solution of $x_n(t)$. Since the latter cannot easily be obtained, a solution of $x(t)$ is found in series form. As shown in reference 1, the coefficients, C_n of this series gives the singularities of

$$\mathcal{L}\{x(t)\} = \lim_{n \to \infty} \mathcal{L}\{x_n(t)\}$$

which may not be poles. These singularities form a convex singularity hull. If this hull is in the left half-plane of s, the Laplace complex variable, the system is stable; otherwise, it is not.

For absolute stability, the convex singularity hull must be in the left half plane for all inputs. To do this practically white noise must be applied to the system, as described in the paper, and the average position of the convex singularity hull must be noted. If this average position is in the left-plane the system is absolutely stable.

Let us now take up the second point raised by Kazda and Bilal namely, "under what conditions would partitioning at points other than the highest derivative be desirable." There are a number of instances where this is desirable. When there is interest in displaying certain harmonics of the system partitioning might be utilized so as to display these in the solution expansion as a linear combination. Another place where partitioning at a point other than the highest derivative is important is the case where the recurrence relations can be made to terminate, or at least be made simpler.[1] There are two important subcases where this simplication takes place. One pertains to the case of a given expansion of the auxiliary function. In this case the partition point can be varied by trial to see if the recurrence relations become simpler. The other subcase pertains to changing the form of expansion of the auxiliary function. For example, if the partition is to be made so that the highest derivative say the k-th and the $(k\text{-}1)$st remain in the partition, the Dirichlet series might be used as an expansion of the auxiliary function. This wou give the simplest moment function, and t solution of the system would be anot Dirichlet series.

Regarding the third point of Kazda a Bilal's discussion, the auxiliary function the Dirichlet series would be selected. T point of partition would depend on how c wished to display the result. If the resul to show functional dependence on ω_0 in ter of harmonics, the partition would be ma after the second linear term namely ω For other purposes it might be made af $\ddot{x}$. In either case, as shown above, characteristic exponent would cont information relative to the frequency oscillation. Determining the amplitude somewhat more complicated if the se does not terminate. What might be d here is the rewriting of the terms of series as trigonometric functions th utilizing trigonometric identities to disp amplitude and phase information.

References

1. See reference 2 of the paper.
2. See reference 7 of the paper.
3. See reference 8 of the paper.
4. The Taylor Series (book), P. Dienes. ford University Press, New York, N. Y., 1931, 89–90.
5. Reply by A. A. Wolf to Bohn's note on Equivalence of the Taylor-Cauchy Laurent-Cauchy Transform Analysis v Conventional Methods, *Proceedings*, Institut Radio Engineers, New York, N. Y., vol. 49, n Jan. 1961, pp. 358–61.
6. See reference 3 of the paper.

Feedback Compensation: A Design Technique

G. J. THALER, MEMBER AIEE
J. D. BRONZINO, STUDENT MEMBER AIEE
D. E. KIRK, STUDENT MEMBER AIEE

THE DESIGN OF feedback compensation for feedback control systems is a subject which is treated in the technical literature as an undertaking incidental to the design of specific systems.

Most textbooks are concerned primarily with cascade compensation; the few[1–3] that devote reasonable space to feedback compensation treat special cases or reduce minor loops, one at a time. None provide a systematic approach which has general applicability. None of the suggested techniques give much guidance as to suitability of a selected compensation path or compensator. Excessive labor in many applications is expended in repeated trial-and-error solutions.

The technique suggested herein is generally applicable to systems of any order, and to nonlinear systems under certain conditions. It permits design of feedback compensators by following well-known and well-practiced techniques of cascade compensation, and offers some insight into whether selected paths and compensators are appropriate. Moreover, the labor involved is no greater than that required for cascade compensation design.

Fundamental Principles

Numerous specific procedures apply to feedback control system design. In general, however, after the selection of what may be called "unalterable components," the system gain, or type number, or both, are fixed to satisfy specifications of static and steady-state accuracy. A stability check follows, and a need for compensation is usually indicated. When feedback compensation is to be used, a suitable path and compensator must be chosen.

Choice of a path is restricted by availability of signal pick-off points, signal feed-in (summing) points, and suitable and accurate pick-off and me uring devices. For simpler syste path selection is relatively easy; complex systems, theoretical infor tion is limited. Some work[4] has b done, but much more is needed.

The suitability of a compensator, h ever, is more readily judged. Firsts must not affect the low-frequency gai the forward path. When this requ ment is applied, a feedback path, utili gain adjustment only, alters the forw gain at all frequencies, thus affecting steady-state accuracy of a type-O sys in response to a step input or a load d turbance, as well as the steady-state ac racy of a type-1 system in response t ramp input, load torque, etc. First rivative feedback affects the velocity error of a type-1 system following a ra

Paper 61-752, recommended by the AIEE F back Control Systems Committee and approve the AIEE Technical Operations Departmen presentation at the AIEE Summer General M ing, Ithaca, N. Y., June 18–23, 1961. Manus submitted March 16, 1961; made available printing April 18, 1961.

G. J. Thaler, J. D. Bronzino, and D. E. K are with the United States Naval Postgra School, Monterey, Calif.

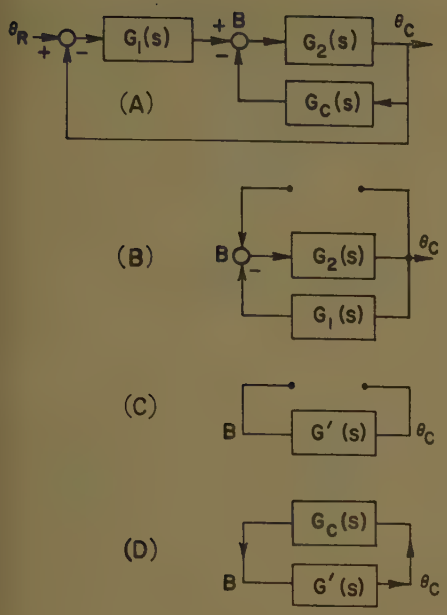


Fig. 1. Fundamental manipulation procedure

A—Typical block diagram
B—Break compensator path and rearrange
C—Reduce to equivalent single block
D—Reconnect compensator

but does not affect the steady-state response to a step or a load disturbance. Second-derivative feedback affects acceleration-lag error of a type-2 system.

The compensator must, as a second requirement, be capable of stabilizing the system and of adjusting the transient response to meet specifications. Aside from such considerations as physical realizability and economics, this simply means that the compensator must be able to move the roots of the characteristic equation to suitable locations; or that it must adjust phase and gain margins and resonant frequency to acceptable values. This is equivalent to saying the modes of transient oscillation are specified by root locations only. True, zeros of the system function adjust the amplitudes of oscillation modes and thus may alter system response to the extent of making an apparently acceptable root configuration unacceptable. Conversely, a root in an undesirable location may be essentially nullified by a properly located zero. In general, selection of a compensator that will provide acceptable root locations is the reasonable starting point.

Since the technique herein developed adopts this approach, certain adjustments of the block diagram are permissible. If the characteristic equation and its roots are the sole interest, all command and disturbance inputs may be kept constant at zero for sake of convenience. This permits elimination of several summing points which otherwise would be annoying.

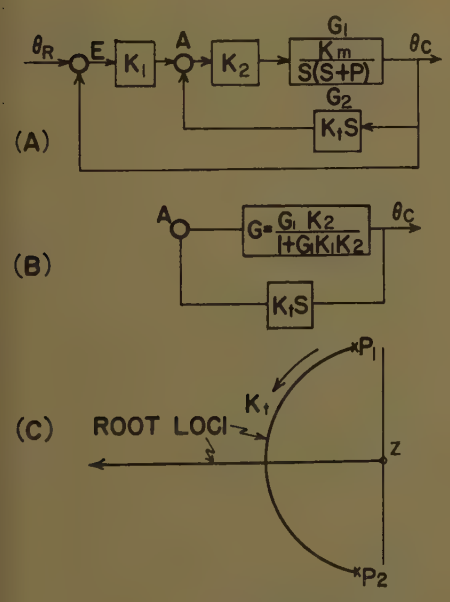


Fig. 2. Tachometer feedback

A—Block diagram
B—Equivalent single loop
C—Root-locus plot

Basic Block Diagram Manipulation

Assume that all gains have been set to satisfy steady-state specifications, and that a feedback compensation path has been selected for trial use. Compensation technique then depends on block-diagram manipulation. The procedure is:

1. Set all input and disturbance signals at zero, and eliminate summation points wherever possible.

2. With compensator path removed, but with its pick-off and feed-in points noted, rearrange block diagram so that feed-in and pick-off points become input and output, respectively. Reduce this diagram to a single equivalent block as in Fig. 1 (A), (B), and (C).

3. Connect compensator to pick-off and feed-in points, thus forming a single-loop block diagram as in Fig. 1(D).

4. Design compensator by using conventional cascade compensation methods.

This technique is applicable to linear systems of any complexity, with the compensator itself either linear or nonlinear. No change in conceptual approach is required if applying this method to certain classes of nonlinear systems, but there will be unavoidable increases in labor.

Second-Order Systems

Tachometer Feedback Compensation

For compensating lightly damped instrument servos, tachometer feedback is commonly used as illustrated by Fig. 2(A), where the motor-load unit is assumed to be of second order. Block-diagram manipulation produces Fig. 2(B), for which the root-locus plot is as shown in Fig. 2(C). Adjustable gain is K_t, the tachometer constant, on this plot. An increase in K_t drives the roots from the poles, increasing the damping but leaving ω_n unchanged.

The ordinate and abscissa of this root-locus plot may be nondimensionalized to form a universal nomograph for tachometer compensation design[5] of static positioning systems. Instead of resorting to a nomogram, however, Appendix I shows that

$$K_t=\frac{2\zeta\omega_n-p}{K_2K_m} \qquad (1)$$

where K_2, K_m, p, and ω_n are defined by the uncompensated system, and ζ is specified by the desired root location.

Two ways of using this equation are apparent. Given the values for motor pole and gain, and with the threshold error specified, the required loop gain $K_1K_2K_m$ is determined. If the gain division into K_1 and K_2K_m is made arbitrarily, then K_t is evaluated from the equation and tachometer circuit designed.

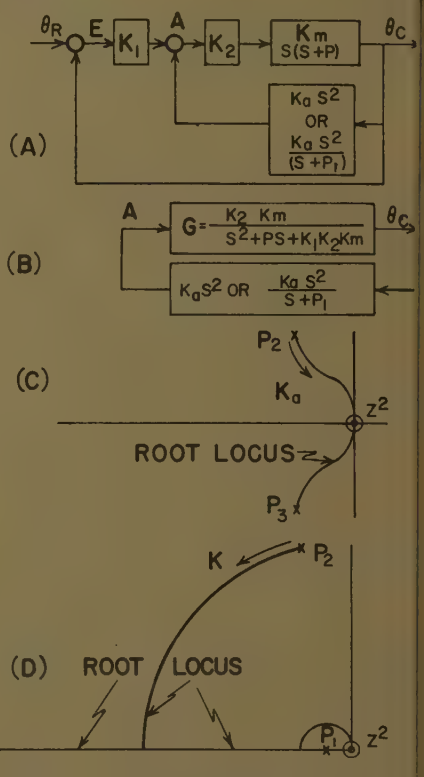


Fig. 3. Acceleration feedback; exact and approximate

A—Block diagram
B—Equivalent single loop
C—Root locus for K_as^2, or for $K_as^2/(s+p$ if $p_1>>p$
D—Root locus for $K_as^2/(s+p_1)$ if $p_1<<p$

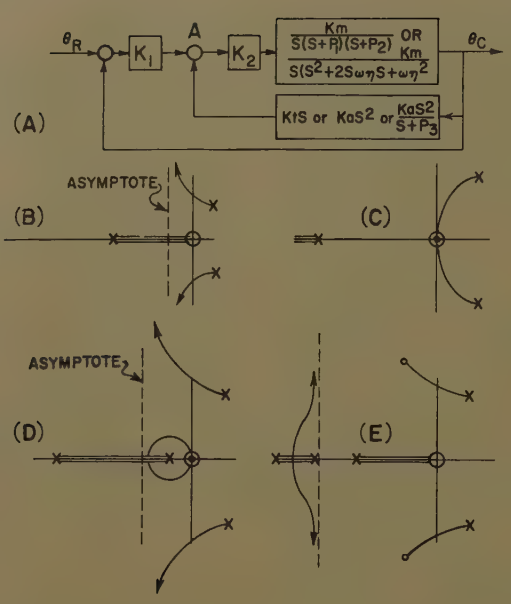


Fig. 4 (left). Third-order systems

A—Block diagram
B—Root locus for tachometer feedback
C—Root locus for acceleration feedback
D—Root locus for approximate acceleration feedback
E—Root locus for feedback through $\frac{K_t s(s^2+\alpha s+\beta)}{(s+p_4)(s+p_5)}$

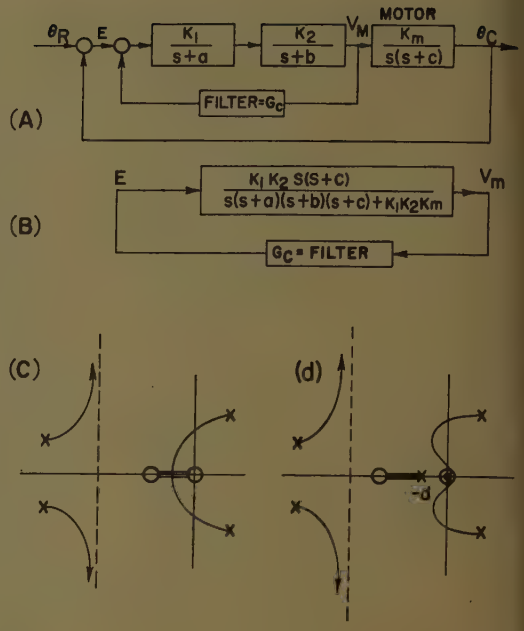


Fig. 5 (right). A fourth-order system

A—Block diagram
B—Equivalent single loop
C—Root locus for $G_c=K$
D—Root locus for $G_c=K_c s/(s+d)$

On the other hand, if a tachometer of known K_t is used, the equation permits evaluation of K_2K_m and thus specifies the gain subdivision which permits use of full K_t, or any selected fraction thereof. For example, if $K_m=200$, $p=23$, and threshold accuracy requires that $K_1K_2K_m=10^5$, then it may be decided arbitrarily that $K_1=100$ and $K_2K_m=10^3$. In this case, for a compensated system with $\zeta=0.7$, the gain must be

$$K_t=\frac{2\zeta\omega_n-p}{K_2K_m}=\frac{2(0.7)10^{5/2}-23}{10^3}=0.42$$

On the other hand, if K_t is 0.32 volt per radian per second, and this full value is used, then the necessary gain subdivision in the forward path is found by noting that

$$K_2K_m=\frac{2\zeta\omega_n-p}{K_t}=\frac{2(0.7)10^{5/2}-23}{0.32}=1{,}310$$

$$K_1=\frac{K_1K_2K_m}{K_2K_m}=\frac{10^5}{1{,}310}=76.3$$

$$K_2=\frac{1{,}310}{200}=6.5$$

Acceleration Feedback Compensation

Since tachometer feedback increases lag error in following a ramp, acceleration feedback is often used to damp transient oscillations of second-order systems. The basic block diagram, Fig. 3(A), after manipulation, takes the form of 3(B). Two transfer functions are shown for the compensator,

K_as^2

for pure acceleration feedback, and

$K_as^2/(s+p_1)$

for the commonly used tachometer followed by a lead network.

Note that approximate acceleration feedback (using the filter) actually provides much better control of dynamic performance than pure acceleration feedback, if the pole is suitably located. Fig. 3(C) shows the root locus for pure acceleration feedback. Damping ratio ζ cannot be improved, although decreased bandwidth is possible. If the tachometer-filter combination has the pole far out on the negative real axis, little improvement is available. But, with the pole near the origin, the root locus, as seen in Fig. 3(D), permits improved damping with large or small ω_n, depending on the amount of feedback gain. Although the latter placement of the pole introduces a real root near the origin for small gain, the compensator pole is a zero of the system function. For a compensator pole less than 0.15 times the motor pole, this zero effectively cancels the residue at the real pole if the feedback gain is small enough to keep the complex roots on the outer semicircle. If the gain is large enough to place complex roots on the inner circle, the real root is large enough to neglect. In either case, second-order response obtains; there is no tail to the step response due to the presence of the real root.

For the case of acceleration feedback using the function

$K_as^2/(s+p_1)$

another simple design nomograph[6] can be constructed which is applicable to most cases of interest. For acceleration feedback, also, an equation is available for computation of feedback gain (see Appendix II):

$$K_a=\frac{1}{K_2K_m}\left[a\omega_{nf}-p-p_1+\frac{aK_1K_2K_m}{\omega_{nf}}+\frac{app_1}{\omega_{nf}}-\frac{p_1K_1K_2K_m(a^2-b^2)}{\omega_{nf}^2}\right] \quad (2)$$

where ω_{nf} is the natural frequency of the final or compensated roots,

$a\omega_{nf}=\zeta_f\omega_{nf}$

is the real part of the root value and the other symbols are defined in Fig. 3. This equation reduces to a simpler form if the desired root is to be located on the outer circular arc of Fig. 3(D)

$$K_a\cong\frac{1}{K_2K_m}\times\left[2a\omega_{nf}-p-p_1(1+a^2-b^2)+\frac{app_1}{\omega_{nf}}\right] \quad (3)$$

and, as an approximation, the last two terms may be dropped, since they are small when ω_{nf} is large, giving

$$K_a\cong\frac{2a\omega_{nf}-p}{K_2K_m}$$

These equations are readily applied. Assume that the system previously considered is to be compensated with approximate acceleration feedback instead of tachometer feedback. Then, for roots on the outer semicircle

$\omega_{nf}=\sqrt{K_1K_2K_m}=\sqrt{10^5}=316$; $K_2K_m=10^3$; $p=23$;

choose $p_1=2.3$; $\zeta=0.7$; therefore

$a=216/\sqrt{2}=223.$

$$K_a=\frac{2a\omega_{nf}-p}{K_2K_m}=\frac{446-23}{10^3}=0.423$$

On the other hand, if the system band-

width is to be reduced, the exact equation must be used, and these values apply

$K_1K_2K_m=10^5$; $K_2K_m=10^3$; $p=23$;

choose $p_1=6$; then for $\zeta=0.7$, $a=b=6$, and $\omega_{nf}=6\sqrt{2}$. It follows that

$$K_a=\frac{1}{10^3}\times$$

$$\left[36\sqrt{2}-23-6+\frac{6\times10^5}{6\times2}+\frac{(6)(23)(6)}{6\times2}-0\right]$$

$$=10^{-3}[36\sqrt{2}-29+0.7\times10^5+0.7(138)]$$

$$\cong 700 \quad (4)$$

Note from this illustration that a good approximation exists when $K_1K_2K_m$ is large and ω_{nf} is on the inner circle

$$K_a\cong aK_1/\omega_{nf} \quad (5)$$

Third-Order Systems

Most control systems are at least of the third order, even if the real root may be large enough to neglect, in which case approximation by second-order-system calculations may be adequate. When the real root is not large, analysis of feedback compensator effects is greatly aided by the technique under discussion. Fig. 4(A) shows the block diagram for a case where the loop-transfer function of the minor loop contains all poles of the forward-transfer function; i.e., the compensation signal is fed back around all poles of the forward path.

Forward-path poles may be either real or complex. For the high gains usually required, complex roots are normally in the right-half plane. Figs. 4(B), (C), and (D) are root-loci sketches for different feedback compensators. Clearly, pure acceleration feedback will not work; tachometer feedback may work if the asymptote is in the left-half plane, but then the real root will probably dominate. Approximate acceleration feedback may also work if the asymptote is in the left-half plane, in which case a pair of complex roots with relatively low ω_n probably can be made to dominate.

None of these compensation schemes is satisfactory for a variety of applications. Much time is saved, therefore, if the manipulation technique is used in a preliminary examination of all suggested compensators to eliminate unsuitable schemes. Fig. 4(E) shows a scheme where the feedback function generates suitably located complex zeros, allowing better placement of the complex roots. Whether the roots can be made dominant or not depends on numerical values involved and on the specific system function obtained.

Higher-Order Systems

This manipulation technique may be applied to linear systems of any order. Choice of compensation path depends on application, but the path need not enclose all forward-path poles. This leads to pole-zero configurations with numerous zeros. The choice of compensator, while restricted by steady-state specifications, may be aided by inspecting the uncompensated root locus.

Fig. 5(A) is the block diagram of a fourth-order positioning system, for which the proposed compensation is attempted by feeding back the motor-armature voltage through a filter. Fig. 5(B) is the manipulated block diagram and Fig. 5(C) shows the root locus for the feedback loop connected with $G_c=K$ but without filter. The connection obviously would stabilize the system, but would disturb the steady-state velocity-lag error. Fig. 5(D) shows the addition of a filter with transfer function $K_ss/(s+d)$, which provides a possible compensator, but may not permit satisfactory root location. The next step is to try a more complicated filter, perhaps with complex zeros. Or, a new compensation path may be indicated.

Nonlinear Systems

If a system is nonlinear and linear feedback compensation is to be used, analysis and design may be accomplished using the describing function representation of the nonlinear component. Block diagram, Fig. 6(A), shows a system wherein the compensator G_c feeds around the nonlinearity G_D. After block-diagram manipulation, the equivalent single loop is as shown in Fig. 6(B), and algebraic manipulation provides the stability relationship

$$G_4(G_c+G_1G_2)=-1/G_D \quad (6)$$

In like manner, Fig. 6(C) shows a system with nonlinear element outside the compensating loop; Fig. 6(D) is the equivalent single-loop diagram, and again the stability relationship is produced by algebraic manipulation

$$-\frac{1}{G_D}=\frac{G_1G_3G_4}{1+G_1G_3G_4} \quad (7)$$

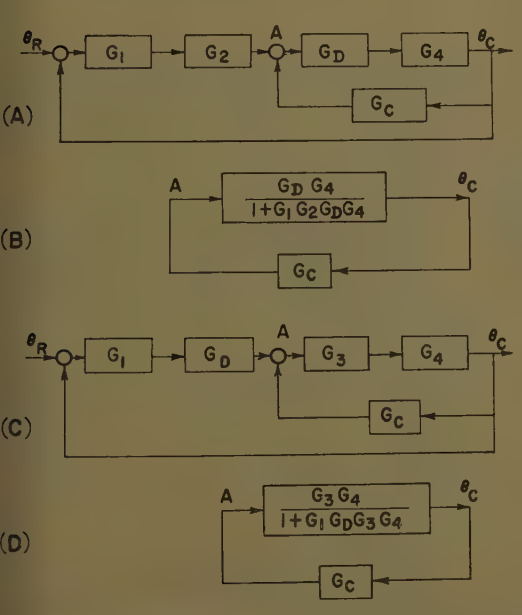


Fig. 6 (left). Linear compensation of a nonlinear system

A—Block diagram, compensation feedback around nonlinearity
B—Equivalent single loop
C—Block diagram, compensation feedback around linear elements only
D—Equivalent single loop

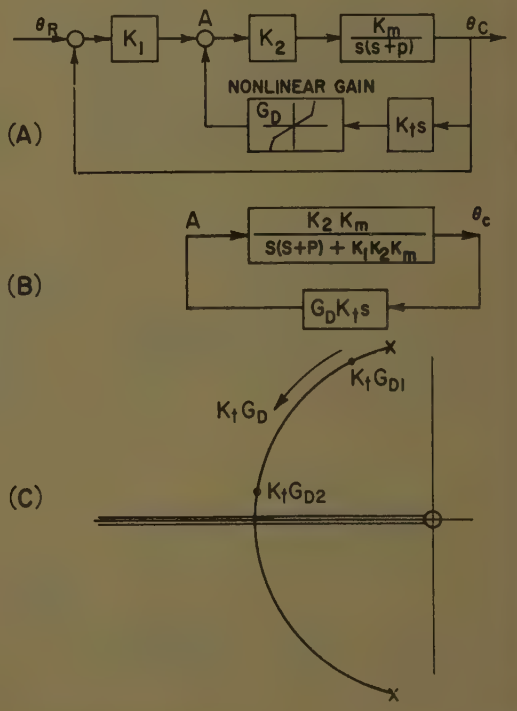


Fig. 7 (right). Nonlinear compensation of a linear system

A—Block diagram
B—Equivalent single loop
C—Root-locus diagram

Equation 6 and equation 7 may be analyzed for stability by Nyquist methods or by the root-locus method if the describing function is amplitude-sensitive only.

If nonlinear compensator G_{Dc} is used with a linear system, then this technique is particularly helpful. Fig. 7(A) is the block diagram of a second-order servo, which is lightly damped to provide fast rise time and small velocity-lag error. If heavy damping is needed for large disturbances, tachometer feedback through a nonlinear amplifier can be used. The nonlinearity is adjusted to give proportionate feedback gain—high for large disturbances; low or zero for small. Using the equivalent single-loop diagram in Fig. 7(B), and the root locus in Fig. 7(C), it is easy to design the nonlinearity with small-disturbance roots at K_tG_{D1}, and with large disturbance roots at K_tG_{D2}.

Conclusions

Design of feedback compensation to control root location, which, in turn, controls stability and dynamic performance, can be reduced to an equivalent problem in cascade compensation by a simple manipulation of the block diagram. In conjunction with this, the root-locus method provides a particularly convenient means of analysis and design. These methods permit derivation of such simple equations as for the tachometer, and also simplify the logical choice of compensation paths and compensators for complex systems.

As some nonlinear systems also may be treated with this technique, nonlinear compensator design thus may be made easier in such cases. Linear compensator design for nonlinear systems is feasible, too, and labor is reduced in the process when considered in comparison with other methods.

Appendix I

Derivation of Tachometer Feedback Relationships

From Fig. 2(B)

$$K_t s\left(\frac{K_2K_m}{s(s+p)+K_1K_2K_m}\right)=-1$$

$$K_t=-\left(\frac{s(s+p)+K_1K_2K_m}{K_2K_m s}\right) \tag{8}$$

Let the desired root be at $s=(-a+jb)\omega_n$, and note that $a^2+b^2=1$. Substitute for s, and consider only the real part of the result, since K_t is a real number

$$K_t=-\left\{\frac{[\omega_n^2(1+a^2-b^2)-ap\omega_n](-a)+b^2(p\omega_n-2a\omega_n^2)}{K_2K_m\omega_n(a^2+b^2)}\right\}$$

$$=\frac{2a^3\omega_n^2+2ab^2\omega_n^2-a^2p\omega_n-b^2p\omega_n}{K_2K_m\omega_n}$$

$$=\frac{2a\omega_n-p}{K_2K_m}=\frac{2\zeta\omega_n-p}{K_2K_m} \tag{1}$$

where $a=\zeta$ since $-a\omega_n=-\zeta\omega_n$ is the real part of the root number.

An alternate and even simpler derivation obtains from the characteristic equation:

$$s^2+(K_2K_mK_t+p)s+K_1K_2K_m=0 \tag{9}$$

and

$$K_2K_mK_t+p\triangleq 2\zeta\omega_n \tag{10}$$

thus

$$K_t=\frac{2\zeta\omega_n-p}{K_2K_m} \tag{1}$$

Appendix II

Derivation of Acceleration Feedback Relationships

From Fig. 3(B)

$$\frac{K_aK_2K_ms^2}{(s+p_1)(s^2+ps+K_1K_2K_m)}=-1$$

$$K_a=-\frac{(s+p_1)(s^2+ps+K_1K_2K_m)}{K_2K_ms^2} \tag{11}$$

Let $s=(-a+jb)\omega_{nf}$, where ω_{nf} is the final or selected natural frequency for the complex roots. Substitute, expand, and consider the real part only

$$K_a=\frac{1}{K_2K_m}\left(a\omega_{nf}-p-p_1+\frac{(K_1K_2K_m+pp_1)a}{(a^2+b^2)\omega_{nf}}-\frac{K_1K_2K_mp_1(a^2-b^2)}{(a^2+b^2)\omega_n{}^2{}_f}\right) \tag{12}$$

but $a^2+b^2=1$, thus

$$K_a=\frac{1}{K_2K_m}\left(a\omega_{nf}-p-p_1+\frac{aK_1K_2K_m}{\omega_{nf}}+\frac{app_1}{\omega_{nf}}-\frac{p_1K_1K_2K_m(a^2-b^2)}{\omega_{nf}{}^2}\right) \tag{2}$$

K_1, K_2, K_m, p, and p_1 are known from the system components, while a, b, and ω_{nf} are specified by selecting the desired root location. If it is on the inner circle of the root locus, then the foregoing equation must be used; if on the outer circular arc, $K_1K_2K_m=\omega_{nf}{}^2$ so that

$$K_a\cong\frac{1}{K_2K_m}\left(a\omega_{nf}-p-p_1+a\omega_{nf}+\frac{app_1}{\omega_{nf}}-p_1(a^2-b^2)\right) \tag{3}$$

which indicates that p_1, $p_1(a^2-b^2)$, and app_1/ω_{nf} are small and may be neglected. Then

$$K_a\cong\frac{1}{K_2K_m}\ (2a\omega_n-p)$$

An alternate derivation of the latter equation is

$$K_aK_2K_m+p+p_1=\Sigma\text{ roots}=2a\omega_{nf}+r_3 \tag{13}$$

from which is taken

$$K_a=\frac{2a\omega_{nf}-p+r_3-p_1}{K_2K_m} \tag{14}$$

For $r_3\cong p_1$, with both p_1 and r_3 small and ω_{nf} large, this readily reduces to the foregoing unnumbered equation. By combining equations 14 and 2, an exact expression for the real root location is given.

$$r_3=-a\omega_{nf}+\frac{aK_1K_2K_m}{\omega_{nf}}+\frac{app_1}{\omega_{nf}}-\frac{p_1K_1K_2K_m(a^2-b^2)}{\omega_n2_f} \tag{15}$$

Combining equations 14 and 3 gives an approximate value

$$r_3\cong p_1\left(\frac{ap}{\omega_{nf}}-a^2+b^2\right) \tag{16}$$

References

1. Introduction to the Design of Servomechanisms (book), J. L. Bower, P. M. Schultheiss. John Wiley & Sons, Inc., New York, N. Y., 1958.

2. Control System Analysis and Synthesis (book), J. J. D'Azzo, C. H. Houpis. McGraw-Hill Book Company, Inc., New York, N. Y., 1960.

3. Elements of Servomechanism Theory (book), G. J. Thaler. McGraw-Hill Book Company, Inc., 1955.

4. An Analytical Technique for the Design of Multi-loop Servomechanism Compensation, J. D. Beecher, A. M. Pride. *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., May 1960.

5. First Derivative Feedback Compensation of Control Systems, D. E. Kirk. *M. S. Thesis*, United States Naval Postgraduate School, Monterey, Calif., June 1961.

6. Compensation of Control Systems Utilizing Acceleration Feedback, J. D. Bronzino. *M. S. Thesis*, United States Naval Postgraduate School, June 1961.

Discussion

Thomas J. Higgins (University of Wisconsin, Madison, Wis.): This paper gives an easily grasped account—from the root-locus approach—of procedure for analyzing control systems whose steady state and transient performance may be improved by inserting a compensating network in the feedback rather than the forward link. No textbooks on basic control theory now in print contain similar content, although such information should be enfolded in every introductory study course utilizing such a text.

The authors have filled a gap in both educational and practical aspects of basic control engineering theory. Their clear exposition, well-buttressed by numerical examples, illustrates the ways in which basic theoretically derived relationships are applied.

J. J. D'Azzo and **C. H. Houpis** (Air Force Institute of Technology, Wright-Patterson

Air Force Base, Ohio): The technique presented in this paper is a definite contribution in the area of feedback compensation.

Some comments in the opening paragraph, however, regarding reference 2, are incorrect. Sections 14-8 and 14-9 of that textbook explain a feedback compensation technique (Method 1) with a systematic approach and limited general usefulness. Some insight also is provided into feedback compensator suitability.

The method presented in the text has two advantages over that given in the paper:

1. Output and input points are maintained.

2. The selected compensation's effect on gain, and therefore on steady-state accuracy, can be determined with greater ease. Both methods are handicapped in that they may require the solution of a third- or higher-order polynomial. Using the technique described in the paper, polynomial complexity depends on the original system's complexity. Combining this system into one transfer function with factored numerator and denominator can be tedious. Polynomial complexity depends on feedback compensator complexity when using the referenced method, and this, likewise is undesirable.

W. C. Schultz (Cornell Aeronautical Laboratory, Inc., Buffalo, N. Y.): The authors apparently have a good scheme for quickly determining the preliminary design of a feedback compensator network. This conclusion is circumscribed by the word "apparent," since the procedure is not clearly outlined. Hence, this question: How does the new technique differ from that proposed earlier by Ross, Warren, and Thaler?[1] The value of both papers might be increased if some statements, and possibly examples, were given on how to use both methods in combination.

A second question concerns implied statements that this method is generally applicable when nonlinearities are encountered. A bit more caution would seem necessary if dealing with nonlinear rather than linear systems. Would the authors please comment on this point?

REFERENCE

1. DESIGN OF SERVO COMPENSATION BASED ON THE ROOT LOCUS APPROACH, E. R. Ross, T. C. Warren, G. J. Thaler. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 79, Sept. 1960, pp. 272–77.

G. J. Thaler, J. D. Bronzino, D. E. Kirk: We thank Professors Higgins, D'Azzo, and Houpis, and Dr. Schultz for their interest and discussions, and the latter for his additional comments in conversation and correspondence, which led to some significant modifications in our wording.

Surprisingly, the concepts in this paper have not been exploited heretofore, even though, as Professor Higgins points out, the material is basic and important. Certainly we shall incorporate it in introductory control courses.

With regard to comments by Professors D'Azzo and Houpis and Dr. Schultz, we feel that somehow an erroneous impression has been conveyed; that is, that our research has uncovered a design procedure which can be specified, step by step. This is not correct, and we are sorry if any misunderstanding has arisen. We deliberately used the word "Technique" in our title. The paper itself is primarily descriptive of block diagram manipulation as a short-cut way of reducing feedback compensation problems to cascade compensation problems, thus permitting application of any well-developed cascade compensation procedure.

The technique does not select the best feedback path or type of compensator (except by trial and error), nor does it propose any specific form of numerical or graphical calculation. In our experience, the trial-and-error analysis (leading to selection of best paths and best compensators) has been most convenient and illuminating when the root-locus method is used. This is a personal preference, however. Bode diagrams will do the same job and may be favored by many designers. With either tool, the first step is to analyze the uncompensated system and determine whether or not compensation is needed, and if so, to what extent.

Results of such analyses are not discarded, but are put to further use in conjunction with our technique. The "equivalent single block," to which the uncompensated system is reduced by block-diagram manipulation, is always the closed-loop system function, multiplied by some quantity (constant or transfer function). Thus, if a root-locus analysis has been used originally, the uncompensated system's roots are the poles of the equivalent block, and the zeros (if any) must be calculated. The compensator block merely represents additional poles and zeros which are arbitrarily placed at selected locations to force the compensated system's roots to lie where we want them. Thus, sketches may be used, and the compensator poles and zeros moved around until the root locus looks satisfactory. This merely establishes the TYPE of compensator; thereafter, numerical design proceeds as usual with standard tools. If the original analysis used Bode diagrams, the "equivalent single block" represents closed-loop frequency response, multiplied by some function, and the resultant Bode plot is easily computed. The compensator block is just a cascaded transfer function, and all Bode diagram design techniques apply.

As to specific comments by Professors D'Azzo and Houpis, the illustration in their textbook deals with pure tachometer feedback around a third-order motor, and there is an amplifier (gain only) between the error detector and the point at which the tachometer signal is added. Their method is to formulate the characteristic equation and rearrange it into the form $1+F(s)=0$ where $F(s)$ is chosen so that a parameter (K_t in this case) is conveniently located in the numerator. The original input and output terminals are retained because the denominator of $F(s)$ is the same as that of the uncompensated loop $G(s)$. In comparison with the tachometer feedback illustration in the paper, the D'Azzo-Houpis manipulation has some advantages. However, our tachometer feedback illustration was merely an example of applying the technique. If the amplifier in the D'Azzo-Houpis system contains a filter, or if the compensator is not pure tachometer feedback, but also contains a cascaded filter, then their approach does not necessarily retain the original input-output points, nor does it isolate the tachometer K_t conveniently. Certainly it is not of assistance in choosing values for the pole and zero in the filter compensator. On the other hand, the technique developed in the paper retains the compensator block as a cascade unit and offers insight into the problem of compensator design.

With regard to Dr. Schultz's first question, this is easily answered in view of the preceding discussion. The Ross-Warren-Thaler procedure complements the technique developed in this paper. After selecting compensator path and type by the technique explained in this paper, the Ross-Warren-Thaler method may take over the task of carrying out the numerical design.

Dr. Schultz's second question is more difficult to answer. First, root-locus methods probably will not be as useful in nonlinear as in linear cases, despite the fact that they do apply nicely to the illustration in the paper. Frequency-response methods are likely to be more effective in general. Second, if the nonlinearity is in the compensation path, or if the compensator itself is nonlinear, the technique applies in the usual manner and we are aware of no restrictions or difficulties except those usually encountered in describing function work. In general, when the nonlinear element is in the uncompensated system and in some path other than the compensator path, then the proposed technique merely serves as a means of deriving the transfer function equations and not much more. In some cases the result may assist in the compensator design, and in others it may not. Many additional explorations could be made in this area.

Can Electric Actuators Meet Missile Requirements?

G. C. NEWTON, JR.
MEMBER AIEE

R. W. RASCHE
NONMEMBER AIEE

ONE OF THE vital components of a missile is the device used to position the thrust directors or control surfaces. In addition to satisfying rather stringent dynamic performance requirements, this device must be highly reliable. Missile designers favor hydraulically operated actuators because of their high dynamic performance capabilities, relatively low weight, and small size.[1] However, the hydraulic actuators have still not reached the desired reliability because of sensitivity to dirt particles and other factors.

A number of other approaches have been considered for missile control surface actuation in addition to hydraulic actuators. Clutch-type servomechanisms have been developed but subsequently abandoned because of unsatisfactory reliability. A number of gas-operated fluid power systems have been considered and some of these are under development.[1] Among these is a gas-operated reaction-jet servomotor.[2] Potentially this device may offer better reliability, although its dynamic performance capabilities do not appear to be too promising.

Another alternative to hydraulic actuators is the electric motor. With the advent of solid-state control devices for electric energy, it should be possible to develop a highly reliable servo system for control surface actuation based upon an electric motor drive.[3] However, up to the present time, electric motors of acceptable sizes and weights have been unable to meet the dynamic performance requirements of these applications.

This paper examines the theoretical feasibility of developing unconventional electric devices that can meet the imposed performance requirements of missile control surface applications. First, the performance parameter that is most definitive in determining the fitness of an electric motor in a high-performance application is identified. This parameter is the torque-squared-to-inertia ratio or power rate of the motor. Next, the power rate requirements of typical missile applications are established and it is shown that conventional motors of reasonable size and weight are unsatisfactory. This leads to an examination of the theoretical limit on the power rate of moving-conductor d-c motors. It is found that it is possible, theoretically, to build an electric actuator that will meet missile requirements. Fortified by this result, a tentative design of a servomotor is set forth. The paper closes with a brief examination of the prospects for a reluctance-type motor in missile applications.

Paper 61-711, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE-AIChE-IRE-ISA Joint Automatic Control Conference, Boulder, Colo., June 28–30, 1961. Manuscript submitted January 13, 1961; made available for printing May 5, 1961.

G. C. Newton, Jr., and R. W. Rasche are both with the Massachusetts Institute of Technology, Cambridge, Mass.

The research of this paper has been supported in part by the Office of Naval Research through support extended to the Electronic Systems Laboratory, Electrical Engineering Department, Massachusetts Institute of Technology under Contract *NOnr-1841(53)*. The authors would like to thank the members of the research department of the United Shoe Machinery Corporation who were consulted on a high-performance servomotor problem.

Power Rate as a Motor Performance Parameter

Fig. 1 shows schematically a servomotor connected to its load through gearing (or linkage). In order to answer the following questions:

1. What is the proper gear ratio R that should be used between the motor and load in order to minimize the size of motor required?
2. What is the minimum motor size that can be used?

certain assumptions must be made concerning the motor, load, and gearing.

The motor's controlling system is assumed to be capable of compensating for dynamic lags in the buildup of motor torque so that dynamic lags do not have to be considered in the determination of the motor size. The motor is assumed to have limits on the maximum torque, T_M, and on the maximum angular velocity, v_M, that it can produce. Furthermore, it is assumed that the peak torque and maximum velocity are not interrelated. That is, the motor is assumed to be capable of producing the peak torque at any velocity including maximum as well as zero velocity. This is not an unreasonable assumption with suitable control, since the torque limit is related to magnetic saturation. In addition to the velocity and torque limits the motor is assumed to be characterized by a moment of inertia, J_M.

Several simplifying assumptions are made with respect to the load and gearing. The load is characterized by a peak torque requirement, T_L, a peak acceleration requirement, a_L, a peak velocity requirement, v_L, and a moment of inertia, J_L. It is assumed that the required load accelerations are independent of the load torque and velocity. Thus, the peak acceleration can be demanded at velocities very close to the peak velocity. It is assumed that the gearing has little kinetic energy in relation to the combined kinetic energies of the motor and load; thus, the inertia of the gearing may be ignored or accounted for by slight modifications of the motor and load inertias. In other respects, such as compliance and backlash, the gearing is assumed to be ideal, which means that it can be characterized by a single parameter, namely, the gear ratio, R_G, which is larger than one when the motor velocity is larger than the load velocity.

Starting with these assumptions it is seen that there are two basic relations which must be fulfilled if the motor is to be capable of driving the load: the motor velocity capability must be equal to or greater than that required by the load, thus,

$$v_M \geq R_G v_L \tag{1}$$

and the motor torque capability must exceed the peak torque required by the load after allowance has been made for the torque required to accelerate the motor itself. In terms of the previously defined symbols this may be written as

$$T_M \geq J_M R_G a_L + \frac{J_L a_L + T_L}{R_G} \tag{2}$$

On the basis of these two relationships, Newton[4] discussed how the required motor size for driving a specified load can always be determined where either one or both of the conditions are effective constraints in establishing the motor size. From the results of this paper it can be shown that the second relation governs motor size when the ratio of the peak

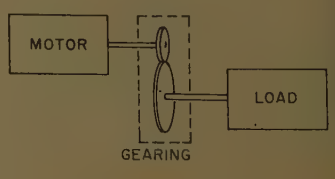


Fig. 1. Motor coupled to load through gearing

Table I. Representative Actuator Requirements for Flap-Controlled Missiles

	Requirement A	Requirement B
Inertia of flap assembly	0.68×10^{-3} kg-m²*	1.09×10^{-3} kg-m²*
Maximum hinge moment	102 newton-meters	51.5 newton-meters
Maximum hinge acceleration	1,100 rad/sec^{-2}†	625 rad/sec^{-2}†
Maximum hinge rate	20 rad/sec^{-1}†	10 rad/sec^{-1}†

* kilogram-meter.
† radians per second.

motor velocity to the peak unloaded motor acceleration equals or exceeds by one half a corresponding ratio for the load. That is, if

$$\frac{v_M J_M}{T_M} \geq \frac{1}{2}\frac{v_L}{a_L} \tag{3}$$

then relation 2 determines the motor size required to drive the load. For electric motors driving missile control surfaces the load acceleration requirements usually are such that condition 3 is fulfilled. In the remainder of this paper this is assumed to be so.

With condition 3 satisfied so that relation 2 governs the motor size, it is known that the motor characteristic which measures its capabilities for driving the load is its torque-squared-to-inertia ratio. This can be seen from relation 2 by finding the gear ratio, R, which minimizes the motor torque required. This optimum gear ratio, found by minimizing the right member of relation 2 with respect to R_G, is given by

$$R_{G(\text{opt})} = \sqrt{\frac{J_L a_L + T_L}{J_M a_L}} \tag{4}$$

This gear ratio may be thought of the one that matches the motor to the load impedance. Substituting this optimum gear ratio into relation 2 shows that the peak torque the motor must have, if the motor is to be capable of driving the load, is given by

$$T_M \geq 2\sqrt{(J_L a_L + T_L) J_M a_L} \tag{5}$$

The minimum permissible peak motor torque results when the equal sign in this relation holds.

In order to obtain a relation without motor parameters on the right side, condition 5 is written as

$$\dot{P}_M \geq 4(J_L a_L + T_L) a_L \tag{6}$$

where

$$\dot{P}_M \triangleq \frac{T_M^2}{J_M} \tag{7}$$

The torque-squared-to-inertia ratio of the motor is the product of the peak motor torque and the peak motor acceleration under no-load conditions. Thus, the torque-squared-to-inertia ratio is a power rate and is therefore indicated by $\dot{P}_M$. Relation 6 shows that the peak power rate of the motor must be four times the peak power rate of the load, or greater, for the motor to be capable of driving the load. Harris,[5] among others, has pointed out the significance of the torque-squared-to-inertia ratio or power rate as a servomotor parameter.

To answer question 2, the peak power rate, given by the right side of relation 6, is computed for the load and multiplied by a factor of 4, which establishes the least value of peak power rate the motor may have in order to drive the load. From tabulated data for lines of motors, the motor which meets the required power rate can be selected. Then the gear ratio can be determined by relation 2 by replacing the inequality sign with an equal sign, and an upper and lower limit on the gear ratio, which may be used in coupling the selected motor to the load, is found. There will be a wider range of possible gear ratios as the margin by which the power rate of the motor exceeds the minimum required power rate increases.

Actuator Requirements and Performance Levels of Available Motors

The limitations of conventional motors used as missile actuator can be demonstrated by considering the two sets of actuator performance requirements, for high-performance flap-controlled missiles, listed in Table I.

To determine the required actuator power rate, the actuator load may be considered to be an inertia and a load torque (hinge moment). The hinge moment of a control surface for a given vehicle velocity generally is an increasing function of the angular deflection, and the maximum moment tends to occur at maximum angle. Because maximum acceleration may be demanded in the vicinity of maximum deflection and maximum angular rate, the hinge moments quoted in Table I are the maximum anticipated values. In both cases shown in Table I, the maximum hinge rate is sufficiently low, in relation to the required acceleration, so that the assumption that there is no actuator speed saturation is justified; that is, relation 3 holds for most, if not all, electric motors.

Using equation 6, the minimum required motor power rates for the two sets of requirements are found to be

$$\dot{P}_{MA} = 4[(0.68\times10^{-3})1{,}100 + 102](1{,}100)\times 10^{-3} = 452 \text{ kw/sec} \tag{8}$$

$$\dot{P}_{MB} = 4[(1.09\times10^{-3})(625) + 51.5](625)\times 10^{-3} = 130.4 \text{ kw/sec} \tag{9}$$

A number of conventional, commercially available electric actuators have been analyzed with respect to power rate at the high overload condition corresponding to a 4% duty cycle. (The power rate increases with decreasing duty cycle.) The results of this analysis are summarized in Figs. 2(A) and 2(B); these are plots of power rate as a function of continuous horsepower rating and motor weight. Requirement A is met by two of the considered actuators, namely, the

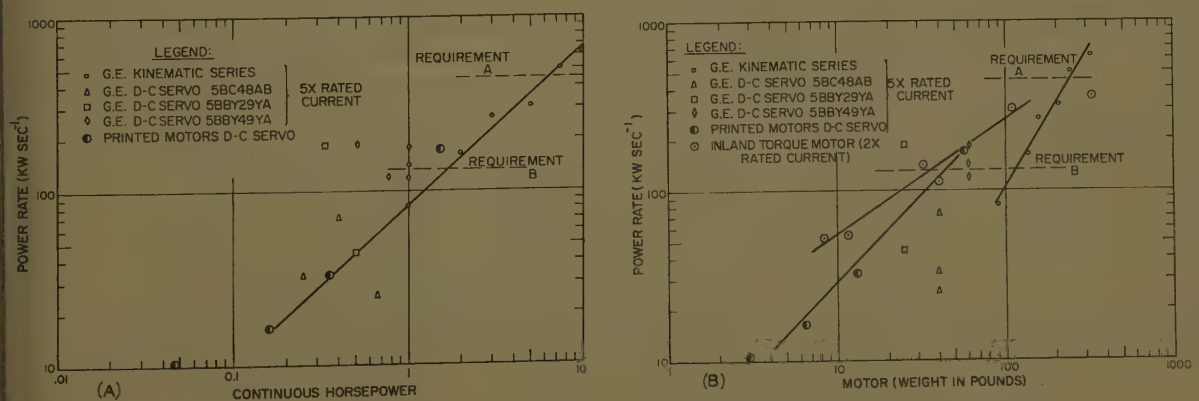


Fig. 2. Power rate versus (A) continuous power rating and (B) weight of typical servomotors

Table II. Comparison of Actuators Meeting Requirements A

Motor	Power Rate, Kw/Sec^{-1}	Weight, Pounds	I^2R Loss, Kw
7½-hp	500	240	7.30
10-hp	622	325	14.60

7½- and 10-hp (horsepower) General Electric kinematic motors which are compared in Table II. For missile applications, actuator weight must be minimized, and thus, it is obvious that the 7½-hp General Electric kinematic motor is the lightest one that will meet the dynamic performance specification.

The conclusion, drawn from this study, is that the motor weight is too great in relation to hydraulic apparatus. The combined weight of the three motors for the required three actuators, in this example, is 720 pounds exclusive of control circuitry and power source. A complete hydraulic system needed to perform this same function weighs approximately 50 pounds including control circuitry and power source. A lesser, though still significant, contrast in weight between electric and hydraulic systems exists for requirement B. This is because the 1/3-hp General Electric servomotor is the lightest motor that can meet the less stringent requirements and this device alone weighs 25 pounds. In view of these findings it is important to determine whether electric devices of suitable weight are, or are not, intrinsically capable of achieving the necessary power rates for missile applications.

This is discussed in the next section for a particular type of motor.

Theoretical Limit on Power Rate of Moving Conductor Actuators

Before rejecting electric motors for missile actuator applications it is desirable to explore the theoretical limit on the power rate of such devices as a function of size. At present it is impossible to discuss such performance limits in general terms; however, for particular types of electric motors limits on power rate can be determined. This section will consider motors with moving conductors of the d-c type. For d-c motors, the two following questions must be answered:

1. How should the motor be proportioned in order to produce the maximum power rate for given size?

2. What is the maximum possible power rate for given size, or conversely, what is the least size the motor may have for a given power rate?

The shell-type motor, shown in Fig. 3, consists of an armature in the form of a hollow shell which revolves around a stationary ferromagnetic core. Field flux through the poles is provided by an external magnetic circuit which is not shown. It is assumed that the magnetic circuit is capable of producing a field flux density, B_f, under the poles. The air gap, which is larger than in conventional motors, is assumed to be sufficiently small so that a substantially uniform flux density exists under the poles. The armature shell is assumed to have a mean resistivity, ρ_r, and a mean density, ρ_m. These values account for conductor insulation and structural materials used to bond the conductors into the shell form.

Utilizing these assumptions, the analysis in the Appendix shows that the power rate of the motor is proportional to the armature I^2R loss, P_R. Specifically

$$\dot{P}_M = \left(\frac{\alpha^2_p}{\alpha_R \alpha_J}\right)\left(\frac{B_{2f}}{\rho_M \rho_R}\right) P_R \qquad (10)$$

where α_p is the fraction of the shell under the field poles, α_R the ratio armature total resistance to the shell resistance, and α_J the ratio of armature total inertia to the shell inertia. Since these quantities cannot be varied appreciably from one motor design to another, it is concluded that the motor power rate is substantially independent of the motor dimensions for a given power loss in the armature.

In answer to the first question using equation 10, the motor proportions have very little influence on the power rate that can be had from a motor of a given size. This result contrasts sharply with the accepted practice in the design of servomotors having a small armature diameter-to-length ratio. The result of equation 10 shows that the important parameter defining the capability of a servomotor, namely, its power rate, is independent of the armature diameter to length ratio and depends only upon the power dissipated in I^2R losses in the armature.

Fig. 4 shows power rate as a function of armature I^2R loss for a number of commercially available d-c motors. The line drawn through the scattered points illustrates that the power rate tends to increase as the 1.105 power of the I^2R loss rather than as the first power as predicted by equation 10. However, in view of the variety of motors plotted in Fig. 4, the agreement is better than one could reasonably expect. It is also interesting to note that 46 seconds^{-1} is the ratio, $\dot{P}_M/P_R$, of the power rate to I^2R loss for a motor of a power rate of 137 kw/sec. This number is a measure of configuration efficiency

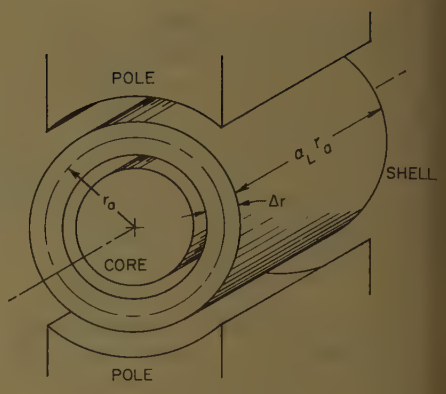


Fig. 3. Shell-type d-c motor

for a motor in producing power rate. For a shell-type motor, equation 10 shows that a ratio, $\dot{P}_M/P_R$, of 724 seconds^{-1} should be obtainable. This figure assumes that $\alpha_p = 2/3$, $\alpha_R = 2$, $\alpha_J = 2$, $B_f = 1.0$ weber meter^{-2}, $\rho_M = 2.71 \times 10^3$ kg-meter^{-3}, and $\rho_R = 5.66 \times 10^{-8}$ ohm-meter. The density corresponds to aluminum on the assumption that aluminum conductors and bonding materials of approximately equal density are used. The resistivity is twice that of aluminum on the assumption that only approximately one half of the shell cross section is conductor. An improvement in power rate capability for a given I^2R loss of more than 15 through configuration improvement certainly appears worth striving for.

The result of equation 10 indicates that there is no limit on the size of a motor for a given power rate, providing ways and means are found to dissipate the armature I^2R losses without excessive temperature rise. However, armature reaction must be considered. In the upper section of Fig. 5, the development of the air gap in the vicinity of one of the poles is shown. The lower portion shows the armature reaction mmf (magnetomotive force) as a function of the angle from the pole center. The peak armature reaction mmf occurs at the edge of the pole. As the dissipated power in the armature is increased, the armature reaction mmf at the pole edge will increase until it is excessively large in relation to the field mmf at this point. Although the armature reaction could be cancelled by pole face windings, this is not feasible in view of the additional I^2R losses.

In the absence of pole face windings the armature reaction mmf must be limited in relation to the field in mmf so that excessive distortion of the resultant flux pattern under the pole is avoided. The ratio of the maximum armature reaction mmf, $F_{a(\max)}$, in Fig. 5, to the field mmf under the pole will be designated as

This ratio is considered to be a de- specification and will have values ging from 1/2 to 1 for motors with shell- e armatures under consideration. In analysis in the Appendix it is shown t the constraint on the armature re- on mmf is equivalent to a constraint he armature shell thickness, Δr. This straint is

$$\geq \left(\frac{\pi \alpha_p^2 \mu_0^2}{2\alpha_R \alpha_L \rho_R n^2{}_p \alpha_a{}^2 \alpha_g{}^2 B_f{}^2} \right) P_{R(\max)} \quad (11)$$

re, μ_0 is the permeability of free space; s the ratio of air gap length to armature l thickness; n_p is the number of poles; l α_L is the ratio of the armature shell gth to its radius. All of the other ibols have been previously defined. us, by relation 11, a lower limit is ced on the thickness of the armature ll, for a specified armature I^2R loss, max), and a specified limitation, α_r, is posed on armature reaction.

Another constraint on the dimensions of armature shell is imposed by cooling siderations. The armature shell must re a sufficiently large surface to trans- the I^2R losses to the cooling medium hin the permissible temperature rise, ax. It is assumed that cooling occurs n the external surface of the armature y. The maximum permitted I^2R loss st be equal to or greater than the aver- I^2R loss that actually occurs. Thus, following equation can be written

$$\alpha_L \alpha_s 2\pi r^2{}_a \Delta\theta_{(\max)} \geq \alpha_d P_{R(\max)} \quad (12)$$

ere U is the over-all heat transfer co- cient; α_s is the ratio of the total ex- nal armature surface to the shell sur- e; $\Delta\theta_{(\max)}$ is the maximum permissible mperature rise of the armature shell ative to the cooling medium; r_a is the an radius of the armature shell; and is the fractional duty cycle. This uation can be written as a constraint on armature radius as follows:

$$\geq \sqrt{\frac{\alpha_d P_{R(\max)}}{2\pi U \alpha_L \alpha_s \Delta\theta_{(\max)}}} \quad (13)$$

uations 11 and 13 establish the dimen- sions of the armature shell except for the ratio α_L, of the length to the radius. There may be some advantage, in equation 11, in making this ratio large in order to reduce the armature shell thickness and thus reduce the length of the air gap. A small air gap involves a small amount of permanent magnet material for the field (or field ampere-turns if electromagnets are used to establish the field flux). However, increasing the length of the armature, in relation to its radius, requires a longer field structure, and offsets, at least partially, the weight saving obtained through the narrower air gap. An exact evaluation of the optimum ratio of length to radius for the armature requires an analysis of the field design problem, which cannot be treated in this paper due to space restrictions. It will suffice to say that commonly used length-to-radius ratios in the vicinity of 2 probably are not too far from optimum.

Thus, the procedure for designing electric motors for missile actuator use is now clear. First, the actuator requirements are analyzed to arrive at a requirement on the motor power rate (torque-squared-to-inertia ratio) as given by relation 6. Next, relation 10 is used to determine the value of armature I^2R loss, P_R, that is necessary to obtain the requisite power rate. Third, using a selected value of armature shell length to radius ratio, α_L, relation 11 yields the lower limit that the armature shell thickness may have without excessive armature reaction. Normally this lower limit is used unless it is prohibited by fabrication difficulties since small air gaps are desirable to minimize the weight of the field structure. Finally, the radius of the armature is established by means of relation 13. Obviously the size of the motor will depend upon the heat transfer coefficient used in relation 11. For missile applications it is necessary to obtain rather large heat transfer coefficients and this may be done by circulating a mist of liquid particles through the air gap. After the basic motor dimensions have been established, the field structure may be designed using standard design procedures. The next section presents a proposed motor design that was developed by this technique.

Design of a Servomotor to Meet Requirement A

In this section, an example of a specially designed shell-type servomotor, which was designed to meet missile actuator requirement A, is given. The parameters of the previous section are assumed. From equation 10 it was shown that $\dot{P}_M/P_R$ is equal to 724 sec^{-1}. The motor power rate needed to meet requirement A is 452 kw/sec^{-1}. If a factor of safety is incorporated and the required power rate is taken as 500 kw/sec^{-1}, the resulting maximum required armature I^2R loss, $P_{R(\max)}$, is

$$P_{R(\max)} = \frac{500}{724} \cong 0.7 \text{ kw} \quad (10A)$$

The minimum allowable shell thickness consistent with the armature reaction constraint now can be found by relation 11. Here, in addition to those parameters previously given:

Ratio of shell length to radius $\alpha_L = 2$
Number of poles $n_p = 2$
Ratio of maximum armature reaction mmf under a pole to the field mmf $\alpha_a = 0.5$
Ratio of air gap length to shell thickness $\alpha_g = 1.25$

Thus, relation 11 gives

$$\Delta r \geq \left[\frac{\pi(0.667)^2(4\pi \times 10^{-7})^2}{(2)(2)(2)(5.66 \times 10^{-6})(2)^2 \times (0.5)^2(1.25)^2(1)^2} \right] 700 \quad (11A)$$

and

$$\Delta r_{(\min)} \cong 2 \times 10^{-3} \text{ meters}$$

The mean shell radius, r_a, is determined by cooling considerations. The following parameters are applicable:

Over-all heat transfer coefficient $U = 340$ watts/m^2C (spray mist cooling assumed)
Duty cycle fraction $\alpha_d = 0.25$

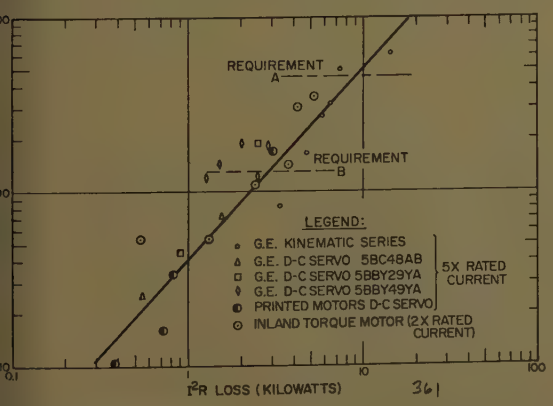


Fig. 4 (left). Power rate versus armature I^2R loss of typical servomotors

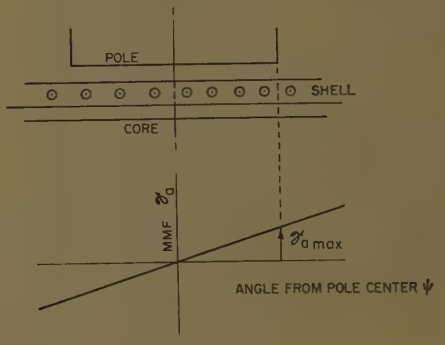


Fig. 5 (right). Armature reaction mmf

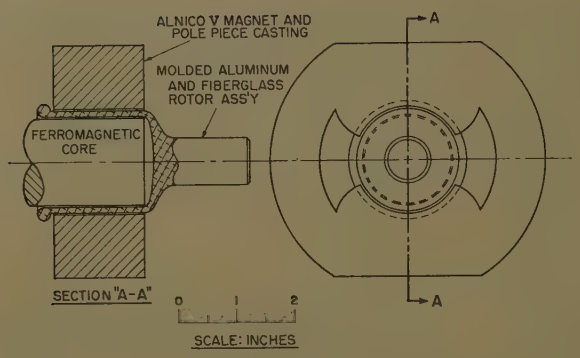


Fig. 6. Sketch of servomotor to meet requirement A

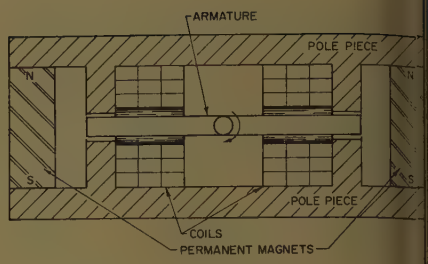


Fig. 7. Reluctance-type motor

Ratio of total armature surface to shell surface $\alpha_s = 2$

Maximum temperature rise $\Delta\theta_{max} = 50$ degrees centigrade

It should be noted that a 25% duty cycle, α_d is used. This value greatly exceeds the 4% that was used in the comparison of conventional motors. With the indicated parameters, relation 13 gives

$$r_a \geq \sqrt{\frac{(0.25)(700)}{2\pi(340)(2)(2)(50)}} \qquad (13A)$$

and

$$r_{a(min)} \cong 0.02 \text{ meter}$$

Thus, the general motor configuration has been determined since shell length, air gap, and pole face width are all related to either Δr_a or r_a by the various design ratios.

A sketch of the rotor and pole piece of the motor is shown in Fig. 6. It can be shown, by conventional analysis, that the permanent magnet configuration will provide the required 1 weber-per-square-meter flux density (10,000 gauss). With 200 armature conductors the brush current will be approximately 31.2 amperes at maximum torque.

The motor, illustrated in Fig. 6, has the following characteristics:

Shell moment of inertia $J_m = 2.18 \times 10^{-5}$ kg-m^2

Maximum motor torque $T_m = 3.32$ newton-meters

The resulting power rate is the required value of 505 kw/sec^{-1}. Therefore, the motor design is consistent with the load requirements. The ferromagnetic material of the design weighs only 7 pounds and the estimated weight of the motor, exclusive of cooling equipment, is less than 11 pounds. In a 3-actuator system, the weight of the cooling system and control circuitry should be less than the weight of the motors alone. Thus, an over-all system weight of 50 to 60 pounds should be achievable. This is exclusive of the increment in power supply weight needed to supply the required electric power. Since the hydraulic system needed to meet these requirements weighs about 50 pounds it is evident that electric actuators may be feasible in this application.

It should be noted that if a 2.5% duty cycle were assumed, no spray mist cooling would be required for the proposed design. Also, requirement A is representative of a rather severe set of actautor requirements. With a reduction in duty cycle or performance requirements, the weight of an electric actuator system can be further reduced and its competitive position in relation to hydraulic systems can thereby be improved.

Prospects for Reluctance-Type Motors

In the preceding section a specially designed, moving conductor, d-c motor was shown to have reasonably good prospects for meeting the power rate and weight requirements that are typical of missile actuator applications. The question arises whether this is the only type of motor that shows such promise. A partial answer to this question is supplied by briefly examining reluctance-type motors in which torque or force is developed on the moving element because of a changing reluctance of an air gap. A relay is an elementary example. Fig. 7 shows a more sophisticated type of reluctance motor. An analysis of the theoretical limit on the power rate of this type of motor, similar to that conducted for moving conductor devices, shows that the ratio of power rate to I^2R loss ranges from 10,000 to 50,000 inverse seconds. These figures are substantially larger than for moving conductor devices which range from 100 to 1,000 inverse seconds.

In order to lend credence to these figures, data are cited in Table III for a commercially available torque motor for actuating hydraulic control valves that utilizes the configuration of Fig. 7.

Unfortunately, the favorable power rate to I^2R ratios for reluctance-type motors are obtained only for small air gaps, and therefore, for limited linear or angular motion. To obtain continuou rotation it is necessary to employ a plura ity of motors of this type in conjunctio with a mechanism for converting oscill tory motion into continuous rotatio This mechanism would be analogou to connecting rod and crank mechanism used in reciprocating engines. If th inertia introduced by the coupling mechan ism is negligible in relation to the bas motor inertias, it can be shown that fo two motors moving sinusoidally 90 de grees out of phase, the resultant torque squared-to-inertia ratio or power rate a the output shaft will be equal to tha characterizing the individual motors Thus, assuming that suitable couplin mechanisms or linkage can be found, th reluctance-type motor should be able t provide the basis for an actuator of powe rate capability that would be quite satis factory for missile applications.

Another disadvantage of the reluctance type motor is the necessity of supply ing the individual motor elements wit alternating currents of adjustable fre quency. In effect, a plurality of suc motor elements driving a common shaf in the manner described constitutes synchronous motor. For a typical missil application, the frequencies of the sup plied currents will range from zero t several hundred cycles per second. Th complication of a variable-frequency source that is capable of supplying con trolled currents against back electro motive forces of amplitude that ar proportional to frequency, constitutes serious handicap for this type of motor.

Table III. Data on Torque Motor

Manufacturer	Midwestern Geophysica Laboratory
Model no.	9
Moment of inertia J_M	64×10^{-6} inch-pound-sec
Peak torque T_M	8.33 inch-pounds
Power rate $\dot{P}_M$	132 kw sec^{-1}
Peak current i_c	40 milliamperes
Coil resistance R_c	3,400 ohms
Coil inductance L_c	7 henrys
I^2R loss P_R	5.44 watts
Ratio $\frac{\dot{P}_M}{P_R}$	24.3×10^3 sec^{-1}
Weight	1.2 pounds

Considering both the complications of ne mechanism required to convert scillatory to rotary motion and the re-uired variable-frequency current source, he prospects for reluctance-type motors or missile actuator applications appear to e questionable, in spite of the large ower rates. However, the harmonic-rive concept[6] may provide an answer o the problem of converting oscillatory o rotary motion, and therefore, in view f their large theoretical power rates, esearch will continue on actuators using eluctance-type motors.

Conclusions

This paper has shown that the torque-squared-to-inertia ratio or power rate f a motor is the most important char-cteristic for determining its applicability s a missile control surface actuator. Examination of typical missile control urface requirements shows that the equired power rates range form 100 to 500 kw per second.

An examination of conventional motors f both the ordinary and printed circuit ariety has shown that they are incapable f meeting the power rates required in nissile applications within sizes and weights that are useful. However, an nalysis of the theoretical limit on the ower rates that can be achieved by mov-ng conductor d-c motors shows that the ower rate tends to be proportional to the I^2R losses in the armature and independ-nt of the motor size. The dimensions f the armature tend to be fixed by arma-ure reaction and cooling considerations. Using these results, a specially designed l-c motor of reasonable weight and size as shown to be capable of meeting the mposed power rate requirements of a ather difficult missile application. Fi-ally, another type of electric actuator in he form of a reluctance-type motor has een examined. Although this type of notor is capable of rather extreme power ates, it suffers from two serious dis-dvantages: the need for special mechan-sms for converting oscillatory motion nto rotation, and the need for a power upply that is adjustable in both frequency nd current magnitude.

This paper has shown that it is theo-etically feasible to develop electric actua-ors for missile applications insofar as ynamic performance and size and weight equirements are concerned. The over-ll systems problem including electric nergy sources and controls has not been discussed, but this must be considered in each application before deciding whether or not to use electric actuators.

Appendix

The derivation of equation 8 is as follows. The basic expression for the torque developed by the armature is

$$\mathbf{T}_M = \int dv[\mathbf{r} \times (\mathbf{j} \times \mathbf{B}_f)] \tag{14}$$

where $\mathbf{j}$ is the current density in the armature shell and $\mathbf{r}$ is the radius vector to the volume element, dv, from a point on the axis of rotation. The other symbols have been previously defined; all symbols are defined in the nomenclature. The integration is to be carried out over all of the armature volume. Using the previously defined nomenclature the magnitude of the torque turns out to be

$$T_M = 2\pi\alpha_p\alpha_L r_a^3 \Delta r j B_f \tag{15}$$

The moment of inertia of the armature, assuming that there is a thin shell, is given by

$$J_M = 2\pi\alpha_J\alpha_L\rho_m r_a^4 \Delta r \tag{16}$$

so that the torque-squared-to-inertia ratio is

$$\frac{T_M^2}{J_M} = \frac{2\pi\alpha_p^2\alpha_L r_a^2 \Delta r j^2 B_f^2}{\alpha_J \rho_m} \tag{17}$$

The power dissipated in the armature is given by the volume integral

$$P_R = \int dv(j^2\rho_R) \tag{18}$$

For the shell-type armature under consideration the integration yields

$$P_R = 2\pi\alpha_R\alpha_L\rho_R r_a^2 \Delta r j^2 \tag{19}$$

Using this equation to eliminate j from equation 17 yields equation 10 since the torque-squared-to-inertia ratio is defined as the power rate $\dot{P}_M$.

The constraint on the armature shell thickness Δr given by relation 11 is derived as follows. Referring to Fig. 5 the basic armature reaction constraint is

$$\mathcal{F}_{a(\max)} \geq \alpha_a \mathcal{F}_f \tag{20}$$

which states that the maximum armature reaction mmf, $\mathcal{F}_{a(\max)}$, under a pole shall be equal to or less than the field mmf, $\mathcal{F}_f$, multiplied by a specified factor α_a. In terms of the field flux density, B_f, and the armature shell thickness this expression becomes

$$\mathcal{F}_{a(\max)} \leq \alpha_a \frac{B_f}{\mu_o} \alpha_g \Delta r \tag{21}$$

since the air gap is assumed to be proportional to Δr. From Fig. 5 it is evident that the armature reaction mmf at the edge of the pole is given by

$$\mathcal{F}_{a(\max)} = \frac{\pi\alpha_p r_a \Delta r j_{(\max)}}{n_p} \tag{22}$$

(See page 312 for discussion)

Squaring both sides and using equation 19 to eliminate j yields

$$\mathcal{F}_{a^2(\max)} = \frac{\pi\alpha_p^2 \Delta r P_{R(\max)}}{2n_p^2\alpha_R\alpha_L\rho_R} \tag{23}$$

Squaring both sides from equation 21 and using the value of $\mathcal{F}_{a(\max)}$ given by equation 23 yields equation 11 upon rearrangement.

Nomenclature

a_L = peak load acceleration
R_G = ratio of gearing
v_L = peak load velocity
v_M = peak motor velocity
J_L = moment of inertia of load
J_M = moment of inertia of motor
M_M = mass of motor
$\dot{P}_M$ = peak power rate of motor
T_L = peak torque of load
T_M = peak torque of motor
$P_R = I^2R$ loss in armature
r_a = mean radius of armature shell
ρ_R = mean resistivity of armature shell
ρ_m = mean density of armature shell
B_f = mean field flux density under pole
α_p = fraction of shell under poles
α_J = ratio of total inertia to that of armature shell
Δr = thickness of armature shell
α_R = ratio of armature total resistance to that of shell
n_p = number of poles
α_R = ratio of shell length to radius
μ_o = permeability of free space
α_a = ratio of maximum allowable armature reaction mmf under pole to field mmf under pole
α_g = ratio of air gap length to armature shell thickness
$\Delta\theta$ = temperature rise of armature shell
α_s = ratio of total external armature surface to shell surface
α_d = fractional duty cycle
U = over-all heat transfer coefficient
$\mathcal{F}_a$ = armature reaction mmf
$\mathbf{j}$ = current density in armature shell
$\mathcal{F}_f$ = field mmf under pole

References

1. Fluid Power Control (book), **John F. Blackburn, Gerhard Reethof, J. Lowen Shearer**, editors. Technology Press, Cambridge, Mass.; also John Wiley & Sons, Inc., New York, N. Y., 1960.

2. Development of a Gas-Operated Reaction-Jet Servomotor, **R. S. Scher.** *Proceedings*, British Institution of Mechanical Engineers, "Symposium on Automatic Control," London, England, 1960.

3. Application of Silicon-Controlled Rectifiers in a Transistorized High-Response D-C Servo System, **Clarence Cantor.** *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 80, Mar. 1961, pp. 7–12.

4. What Size Motor for Proper Operation of Servomechanism, **G. C. Newton.** *Machine Design*, Cleveland, Ohio, vol. 22, Nov. 1950, pp. 125–30.

5. A Comparison of Two Basic Servomechanism Types, **Herbert Harris.** *AIEE Transactions*, vol. 66, 1947, pp. 83–93.

6. Breakthrough in Mechanical Drive Design: The Harmonic Drive, **C. W. Musser.** *Machine Design*, vol. 32, Apr. 14, 1960, pp. 160–7.

Discussion

N. F. Tsang (University of Arkansas, Fayetteville, Ark.): The authors have given an interesting approach to the solution of high performance actuators. In the specially designed shell-type servomotor it may be noted that a very important design parameter is B_f, the air gap flux density. The allowable amount of total ampere conductors of the armature is proportional to the air gap magnetomotive force, which is in turn proportional to B_f. Hence the torque will be proportional to B_f^2, and $\dot{P}_m$ to B_f^4. A 10% reduction in B_f would mean 34% in $\dot{P}_m$, the performance factor.

It would be enlightening if the authors would give a more detailed analysis of the magnetic circuit of the specially designed motor, stressing the following points:

1. Shape and magnetization of the permanent magnet.
2. Leakage fluxes.
3. Flux reduction due to field distortion.

It seems somewhat doubtful that the assumed B_f of 1 weber-meter^2 could be attained with Alnico V in the design shown. Even though it cannot be realized, some other material, such as Ticonal or Alcomax, may be considered, or the magnetic circuit reshaped, or both.

G. C. Newton, Jr., R. W. Rasche: The shape and size of the permanent magnet shown is intended to represent a minimum size permanent magnet for the device proposed. We recognize that some pole shaping or other means of field compensation or concentration might be required. It is felt, however, that such changes would not result in any appreciable size or weight change. Since the application considered in the paper was an extremely stringent one, the conclusions of the paper would not be affected by minor dimensional changes in the device proposed, or for that matter, by an appreciable degradation of the performance characteristics of the actuator.

The magnetic circuit analysis is based on the assumption that the mmf producing the air gap flux is developed primarily in the side legs of the magnet. The additional magnetic material will only increase air gap flux. The cross-sectional area of each leg is 0.6 A_g where A_g is the cross-sectional area of the air gap under the pole face. A leakage factor of 1.2 is assumed. Fringing is neglected.

Using

$$B_m = \frac{fB_f A_g}{A_m} \tag{24}$$

where B_m is magnet flux density in side legs, f the leakage factor, A_g the cross-sectional area of air gap under pole face, and A_m the cross-section area of magnet side legs, and realizing that, by design,

$$A_m = 1.2A_g \tag{25}$$

The required flux density is

$$B_m = \frac{(1.2)(1.0)(A_g)}{1.2(A_g)} = 1.0 \text{ weber-meter}^{-2} \tag{26}$$

Reference to B-H curves for Alnico V shows 1.0 weber-meter^{-2} corresponds to H_m equal to 35,000 ampere turns per meter^{-1}.

The total gap width is

$$2\alpha_g \Delta r = 0.005 \text{ m} \tag{27}$$

The required length of magnet is

$$L_m = \frac{B_f 2\alpha_g \Delta r \times 10^7}{4\pi H_m} \tag{28}$$

or

$$L_m = \frac{(1.0)(.005)\times 10^7}{4\pi(35{,}000)} = 0.11 \text{ m} \tag{29}$$

The effective length of each side leg is at least 0.11 m depending on how much of the pole pieces are considered. Thus it is felt that the resultant gap flux density will sufficiently exceed 1.0 weber-meter^{-2} to allow for distortion effects. Additional material in the side legs could improve this figure by 25%.

It should be noted that equation 11 is the result of an armature reaction constraint. This constraint can be given any value, but it is our opinion that a value of α_a of 0.5 will result in sufficiently low distortion to assure satisfactory operation.

The suggestions of Mr. Tsang as to the use of Ticonal or Alcomax are valid and are much appreciated.

61-204 Surge Comparison Testing of D-C Armature Windings.....Scheda . . . 537
61-143 Collector Ring and Brush Wear.....Herder, Kerber . . . 543
61-762 Basic Constants of General Induction Machine.....Graybeal . . . 548
61-807 Repeater Receiver for the NEAR System.....Cleary . . . 556
61-633 32-Step Voltage-Regulator Performance.....Gangel, Green . . . 559
61-760 Self-Compensated High-Current Homopolar Generator....Das Gupta . . . 567
61-720 Photoelectric Impulse Generation for Demand Metering.....Whipple . . . 573
61-173 High-Frequency Iron Losses in Fractional-Hp Motors.....Shaneman . . . 579
61-740 Field Patterns of Bundle Conductors.....Timascheff . . . 590
61-761 Saturation Harmonics of Polyphase Induction Machines.....Lee . . . 597
61-489 Mechanism of Breakdown of Laboratory Gaps.....Wagner, Hileman . . . 604
61-488 The Lightning Stroke—II.....Wagner, Hileman . . . 622
61-737 Protection of Lead-Sheathed Power Cables.....Trouard, Maier . . . 642
61-238 Trends in H-V Cable Techniques in Great Britain.....Barnes, Sutton . . . 647
60-832 Determination of Economical Distribution Substation Size.....Smith . . . 663
60-833 Economics of Primary Distribution Voltages of 4.16–34.5 Kv.....Smith . . . 670
61-757 Simplified Methods of Calculating Insulation Life.....Whitman . . . 683
61-772 The Cowans Ford Project.....Wray . . . 685

Conference Papers Open for Discussion

Conference papers listed below have been accepted for AIEE Transactions and are now open for written discussion until December 27. Duplicate double-spaced typewritten copies for each discussion should be sent to Edward C. Day, Assistant Secretary for Technical Papers, American Institute of Electrical Engineers, 345 East 47th Street, New York 17, N. Y., on or before December 27.

Preprints may be purchased at 50¢ each to members, $1.00 each to nonmembers, if order is accompanied by remittance or coupons. Please order by number and send remittance to:

AIEE Order Department
345 East 47th Street
New York 17, N. Y.

60-868 Stabilization of Fluid Servomechanisms by the Stability Factor Method.....Fonda
61-85 Servomotor Characteristics by Impulse Testing.....Weed
61-223 High-Accuracy Digital-Analog Solid-State Speed Controller.....Thompson
61-645 Speed Regulation by Digital Methods.....Potts
61-887 Application and Performance of Insulations for 500 C Hypersonic Aircraft Generators.....Balke, Merrifield, Dimond
61-889 Optimizing Simple Circuitry for Reliability and Performance by Failure Mode.....Hanne
61-891 Static Inverter with Neutralization of Harmonics.....Kernick, Roof, Heinrich
61-906 Problems in the Selection and Testing of Nickel Cadmium Batteries for Satellites.....Albrecht
61-911 Solar-Cell Performance with Concentrated Sunlight.....Tallent, Oman
61-913 Voltage Modulation on Aircraft Electric Power Systems as a Function of the Flexibility of the Alternator Drive Shaft.....Flugstad

AIEE PUBLICATIONS

	Member Prices	Nonmember Prices: Basic Prices*†	Nonmember Prices: Extra Postage for Foreign Subscriptions
Electrical Engineering			
Official monthly publication containing articles of broad interest, technical papers, digests, and news sections: Institute Activities, Current Interest, New Products, Industrial Notes, and Trade Literature. Automatically sent to all members and enrolled students in consideration of payment of dues. (Members may not reduce the amount of their dues payment by reason of nonsubscription.) Additional subscriptions are available at the nonmember rates.		annually $12*	$1.00
		Single copies $1.50*	
Bimonthly Publications			
Containing all officially approved technical papers collated with discussion (if any) in three broad fields of subject matter as follows:			
	annually	annually	
Communication and Electronics	$5.00	$8.00*	$0.75
Applications and Industry	$5.00	$8.00*	$0.75
Power Apparatus and Systems	$5.00	$8.00*	$0.75
Each member may subscribe to any one, two, or all three bimonthly publications at the rate of $5.00 each per year. A second subscription to any or all of the bimonthly publications may be obtained at the nonmember rate of $8.00 each per year.			
Single copies may be obtained when available.	$1.50 each	$1.50* each	
AIEE Transactions			
An annual volume in three parts containing all officially approved technical papers with discussions corresponding to six issues of the bimonthly publication of the same name bound in cloth with a stiff cover.			
	annually	annually	
Part I Communication and Electronics	$4.00	$8.00*	$0.75
Part II Applications and Industry	$4.00	$8.00*	$0.75
Part III Power Apparatus and Systems	$4.00	$8.00*	$0.75
Annual subscription to all three parts (beginning with vol. 77 for 1958).	$10.00	$20.00*	$2.25
Annual subscription to any two parts.		$15.00*	$1.50
AIEE Standards			
Listing of Standards, test codes, and reports with prices furnished on request.			
Special Publications			
Committee reports on special subjects, bibliographies, surveys, and papers and discussions of some specialized technical conferences, as announced in ELECTRICAL ENGINEERING.			

*Discount 25% of basic nonmember prices to college and public libraries. Publishers and subscription agencies 15% of basic nonmember prices. For available discounts on Standards and special publications, obtain price lists from Order Department at Headquarters.

†Foreign prices payable New York exchange

Send all orders to:

Order Department
American Institute of Electrical Engineers
345 East 47th Street, New York 17, N. Y.